# Foundations of Therapeutic Inertia: Understanding determinants of physician's decisions

**Dissertation**
**submitted to the**
**Faculty of Business, Economics and Informatics**
**of the University of Zurich**

to obtain the degree of Doktor der Neuroökonomie, Dr. sc.

(corresponds to Doctor of Neuroeconomics, PhD)

presented by

Gustavo Saposnik
from Canada

Approved in September, 2020 at the request of

Prof. Dr. Philippe Tobler
Prof. Dr. Christian Carl Ruff

The Faculty of Business, Economics and Informatics of the University of Zurich hereby authorizes the printing of this dissertation, without indicating an opinion of the views expressed in the work.

Zurich, 16.09.2020

The Chairman of the Doctoral Board: Prof. Dr. Todd Hare

## Acknowledgments

**To my supervisors:**

*"I wish to express my sincere appreciation to my supervisors, Professor Philippe N. Tobler and Professor Christian C. Ruff, who guided and encouraged me to answer hypothesis-driven research questions and to hold the reader's hand with my writing. A special thanks to Philippe for making himself available Saturdays and Sundays to answer questions and edit manuscripts. Without their persistent support, this project would not have been fruitful."*

**To St. Michael's Hospital leadership, colleague neurologists and participants, and the Division of Neurology at the University of Toronto:**

*"I am most grateful for giving me the opportunity to take a sabbatical year to complete my thesis. Without their support, the coverage of my clinical duties, and funding, this project could not have reached its goal."*

**To my supporters:**

*"A very special thank you goes out to Dr. Jorge Mauriño and the Medical Department at Roche (Spain and Canada) for believing that my innovative work would advance the field despite not having tangible benefits, and for providing funding support. I am also grateful to Rosane Nisenbaum and the Li Ka Shing Institute at St. Michael's Hospital for their technical support."*

**To my wonderful family:**

*"Nobody has been more important to me in pursuing this PhD than my family. I wish to acknowledge the unconditional support, great love and generosity of my family, my wife and life-coach, Amy; my three lovely daughters, Naty, Julie, and Florencia who provide unending inspiration; and my parents, Estela and Pedro, true role models who continue to encourage their grown kid accomplish his goals. Each of them kept me going and this work would not have been possible without their emotional and physical support.*

## List of manuscripts

The dissertation is based on the following research articles:

Study 1:
Saposnik G, Redelmeier D, Ruff CC, Tobler PN. Cognitive biases associated with medical decisions: a systematic review. BMC Med Inform Decis Mak. 2016;16(1):138.

Study 2:
Saposnik G, Sempere AP, Prefasi D, Selchen D, Ruff CC, Maurino J, Tobler PN. Decision-making in Multiple Sclerosis: The Role of Aversion to Ambiguity for Therapeutic Inertia among Neurologists (DIScUTIR MS). Front Neurol. 2017;8:65.

Saposnik G, Sempere AP, Raptis R, Prefasi D, Selchen D, Maurino J. Decision making under uncertainty, therapeutic inertia, and physicians' risk preferences in the management of multiple sclerosis (DIScUTIR MS). BMC Neurol. 2016;16:58.

Study 3:
Almusalam N, Oh J, Terzaghi M, Maurino J, Bakdache F, Montoya A, Caceres F, Saposnik G. Comparison of Physician Therapeutic Inertia for Management of Patients With Multiple Sclerosis in Canada, Argentina, Chile, and Spain. JAMA Netw Open. 2019;2(7):e197093.

Study 4:
Saposnik G, Maurino J, Sempere AP, Terzaghi MA, Ruff CC, Mamdani M, Tobler PN, Montalban X. Overcoming Therapeutic Inertia in Multiple Sclerosis Care: A Pilot Randomized Trial Applying the Traffic Light System in Medical Education. Front Neurol. 2017;8:430.

Study 5:
Saposnik G, Mamdani M, Montalban X, Terzaghi M, Silva B, Saladino ML, Tobler PN, Caceres F. Traffic Lights Intervention Reduces Therapeutic Inertia: A Randomized Controlled Trial in Multiple Sclerosis Care. MDM Policy Pract. 2019;4(1):2381468319855642.

Study 6:
Saposnik G, Oh J, Terzaghi MA, Kostyrko P, Vaidyanathan S,Nisenbaum R, Ruff CC, Tobler PN. How physicians handle uncertainty when making live therapeutic decisions?. This manuscript has been submitted and accepted for publication to JAMA Netw Open.

Study 7:
Saposnik G, Oh J, Terzaghi MA, Kostyrko P, Bakdache F, Montoya A, Jaja BNR, Nisenbaum R, Ruff CC, Tobler PN. Emotional expressions associated with therapeutic inertia in multiple sclerosis care. Mult Scler Relat Disord. 2019;34:17-28.

# Table of Contents

# 1. Introduction

Therapeutic decisions are among the most critical tasks made by physicians in every day clinical practice. Making such decisions requires an individualized assessment of the safety and efficacy profiles of different medications, with either imperfect information or uncertainty about the potential outcomes of treatment choices. Furthermore, those decisions are influenced by a physician's education, personality traits, and cognitive biases. Previous studies revealed that cognitive biases affect over 50% of physicians, with overconfidence, risk tolerance, availability bias, and the framing effect the most commonly reported. From a decision neuroscience perspective, my goal was to better understand how physicians make therapeutic decisions. I was specifically interested in the status quo bias (SQ) when applied to therapeutic decisions in the medical field. SQ is defined as the preference for "keeping things the way they presently are". In medicine, SQ occurs when physicians face different therapeutic alternatives, but end up keeping the current treatment. Therapeutic inertia (TI), a common variant of SQ, is characterized by a physician's decision not to initiate or intensify treatment when treatment goals are unmet. In other words, TI occurs when there is evidence of disease progression (e.g.: elevated blood pressure despite antihypertensive treatment, elevated glucose despite taking medications for diabetes) and physicians do not initiate a new medication or switch to a more effective treatment as recommended by best practice guidelines. Such suboptimal decisions lead to poorer clinical outcomes and higher health care costs.

In the present work, I outline five key questions which are addressed in sequence by my empirical studies. I first begin by introducing the concept of cognitive biases and summarizing the results of an exhaustive literature review (Question 1). Then, I

elaborate upon the concept of therapeutic inertia and its associated factors (Questions 2 & 3). Then, I show the results of a newly designed educational intervention to ameliorate TI (Question 4), Finally, I present the results of studies evaluating the association between arousal response and emotional expressions on therapeutic decisions (Question 5).

The focus of my work is on multiple sclerosis (MS), a demyelinating disease that represents the paradigm of complex treatment decisions in chronic and progressive medical conditions (e.g., hypertension, diabetes, high cholesterol). I first summarize the most common factors associated with TI in MS care. Next, I present the rationale for the development of an innovative educational intervention applying the traffic light system strategy (TLS) to ameliorate the effects of TI. The TLS-based educational intervention is designed to help optimize treatment decisions by interrupting the automatic status-quo state, and trigger re-evaluation processes elicited by the universal warning sign of the color red (Ahmed et al., 2020; Laura Enax, Krajbich, & Weber, 2016).

Thereafter I discuss the results from a feasibility and randomized study testing the benefits our successful TLS-educational intervention on reducing TI. I then explore current understanding of the role of autonomic arousal and emotional status in decision-making, and more specifically our limited understanding of the arousal response in live therapeutic decisions (Saposnik et al, 2020 submitted for publication). Finally, I provide an explanation of the possible link between physician's characteristics (e.g. years of experience, expertise in MS care) and the effects and mechanism by which the TLS educational intervention may reduce TI with the

integration of the arousal state in the modulation of therapeutic choices made by expert neurologists (Saposnik et al, 2020 submitted for publication).

The core of this thesis consists of novel work that begins with the development of the concept of therapeutic inertia in MS care, and then examines changes in the arousal states to test the mechanism by which the TLS educational intervention decreased TI among neurologists who are making live treatment decisions.

My discussion is centered on empirical results and theory: I synthetize the contributions of my studies within the broader literature and suggest a novel model to explain the association between physician's characteristics and the effects of an educational intervention on TI, and the mediating effects of autonomic arousal states on the educational intervention-TI link. I conclude with practical implications for neuroscience, medical education, and patient care.

## 1.1 A brief description of Medical Decisions

Decision making is defined as the process of examining possibilities, evaluation of risks and uncertainties, identifying options, comparing them, and choosing a course of action (Elstein & Schwartz, 2002). Medical decision-making is a complex task involving a variety of cognitive processes, including the selection and integration of best research evidence with clinical expertise and patients' values.(Elstein & Schwartz, 2002; Zikmund-Fisher et al., 2010) A previous report highlighted the importance of teaching about uncertainty to medical students.(Flexner, 2002)

Decisions based on erroneous assessments may result in incorrect patient and family expectations, suboptimal advice, and costly medical errors (Croskerry, 2003). For example, medical errors are the third leading cause of death in the United States of America (Makary & Daniel, 2016). Medical errors represent 1.7-6.5% of all hospital admissions causing up to 100,000 avoidable deaths each year in the USA (Bates et al., 1995; Classen, Pestotnik, Evans, & Burke, 1991). Suboptimal medical decisions and medical errors cost the USA healthcare system approximately $20 billion annually (Andel, Davidow, Hollander, & Moreno, 2012). A recent report from the World Health Organization revealed that four out of ten patients in primary and outpatient care are harmed (https://www.who.int/features/factfiles/patient_safety/en/ accessed April 17, 2020).

The ultimate consequences of suboptimal decisions and medical errors include avoidable hospitalizations, medication underuse (i.e.: therapeutic inertia) and overuse, and wasted resources that may lead to the progression of medical conditions, poorer clinical and cognitive outcomes for patients, and higher healthcare costs (Ioannidis & Lau, 2001; OECD).

**The theoretical framework of medical decision-making**

Medicine is an uncertain science, and physician are not infallible. Most (if not every) physician has made or witnessed some form of a mistake (e.g. medical error, near miss) in diagnosis and/or treatment in their career.(Graber, 2013; Makary & Daniel, 2016) But the frequency of those mistakes, and their severity, can be reduced by understanding how physicians make decisions. Despite the great technological advances in medicine over the last century, uncertainty remains a key challenge in all aspects of medical decision-making. The basic structure of decision models used to evaluate diagnostic or treatment strategies are binary (Harrison, Milbers, Hudson, & Bansback, 2017).

In the last 30 years, several theories have been proposed to define how physicians make decisions. Several concepts have emerged to characterize rational decisions, medical decisions, diagnostic and therapeutic errors, and the status-quo related to lack of treatment initiation or escalation (Ioannidis & Lau, 2001; D. Ontaneda, Tallantyre, Kalincik, Planchon, & Evangelou, 2019; Prakash, Sladek, & Schuwirth, 2019; Redelmeier & Shafir, 1995; Summerfield & Tsetsos, 2015). A recent study provides a comprehensive summary of major theories related to models of rationality relevant to medical decisions, including: the application of Bayesian probability and decision analysis based on expected utility theory (EUT), prospect theory, and dual processing theories (DPT) of rational thought, among others (B. Djulbegovic, Elqayam, & Dale, 2018). Although the explanation of each theory exceeds the scope of this thesis, each of these theories has been assessed regarding information components (e.g.: about options, attributes), deliberation components (e.g.: strategies that help individuals deliberate about their choices: value, preference), and outcomes

measurements. A recent review found that: i) few decision-making interventions are explicitly based on a theory or model, and ii) there was no theory that comprehensively integrated all components or informed about the design of educational interventions aimed at making good decisions.(Bekker et al., 1999; G. Elwyn, Stiel, Durand, & Boivin, 2011)

Briefly, Bayesian variants of threshold models use prior information to formalize calculations (Vilares & Kording, 2011). For example, these models may capture the belief that initiating or intensifying treatment at a certain laboratory threshold provides the best outcome (e.g. initiate treatment for the prevention of diabetes with metformin when blood hemoglobin A1C test levels are greater than 5.7%) (Hostalek, Gwilt, & Hildemann, 2015).

Clinicians commonly face a dilemma when making therapeutic decisions based on the probability of a disease, and the harms and benefits associated with a diagnostic test (e.g. radiation associated with a computerized tomography scan- [CT]) or treatment (e.g. side effects associated with antibiotics, cholesterol lowering therapies). As a result, physicians may order a test to decide whether or not initiate treatment, or initiate treatment without ordering a test. Most recent advances in medical decision making include threshold models (Benjamin Djulbegovic, Hozo, Mayrhofer, van den Ende, & Guyatt, 2019), which rely on the expected utility (an economic measure of the desirability of an outcome that combines the likelihood of the outcome and the satisfaction from it) or variants of these models (Caplin & Glimcher, 2014; Hozo, Tsalatsanis, & Djulbegovic, 2018; Tsalatsanis, Hozo, Kumar, & Djulbegovic, 2015). EUT suggests that when choosing between different strategies, the decision maker

should select the strategy that leads to the outcome with the highest expected utility (Note: the expected utility function is represented in the following formula: $EU(A)=\sum o \in O PA(o)U(o)$; where O is the set of outcomes, PA(*o*) is the probability of outcome *o* conditional on A, and U(*o*) is the utility of *o*.

In medicine, an MRI has a higher expected utility for detecting small stroke or brain lesion in MS than computerized tomography, insulin has a higher utility to decrease blood sugar for patients with uncontrolled diabetes than other oral medications. Medical decisions based on the EUT are therefore based on a "threshold"; the probability at which we are indifferent between testing and administering treatment without testing. Models based on the EUT do not take into account physician's intuition and experience (e.g. a CT may be a more practical test despite the lower expected utility value), or the emotions of the patient (e.g. a patient's ability to tolerate having an MRI due to claustrophobia) or the treating physician when making a decision (B. Djulbegovic, Hozo, Beckstead, Tsalatsanis, & Pauker, 2012).

Dual processing theory (DPT) overcomes the limitations and commonly violated principles of EUT when decisions are based on physician's expertise or intuition and contradict the highest expected value. DPT is based on the concept that human decisions are governed by two distinct processes, commonly referred to as system 1 (intuitive) and system 2 (analytical). In brief, system 1 refers to an automatic, unconscious, fast, and effortless (or routine) mechanism to make most common decisions. Conversely, system 2 makes deliberate decisions, which are non-programmed, conscious, usually slow, and deliberate (Tversky & Kahneman, 1974).

Under the DPT framework, it has been suggested that most cognitive biases are attributed to intuitive processes (representing the overuse of system 1), or when system 1 overrides system 2 (Ely, Graber, & Croskerry, 2011; Mamede, van Gog, van den Berge, van Saase, & Schmidt, 2014; van den Berge & Mamede, 2013). In this framework, techniques that enhance system 2 (e.g. my successful educational intervention to overcome therapeutic inertia) could counteract these biases, and thereby improve diagnostic accuracy and decrease the likelihood of suboptimal decisions and medical errors (e.g. therapeutic inertia) (Saposnik, Mamdani, et al., 2019; Saposnik, Maurino, et al., 2017).

Treatment decisions based on DPT suggest that physicians decide to treat when the "threshold probability" at which treatment benefits are greater than treatment harms are exceeded. However, they also incorporate other critical components not considered under the EUT, such as intuitive cognitive processes, emotional aspects, balance between efficacy and side effects of medications, and the experience of decision-makers. Physicians also make automatic decisions when specific criteria are met (e.g. to initiate antibiotics if fever and signs of an ear infection are present by direct visualization in a child with otitis, not requiring a threshold probability). Decisions based on diagnostic tests are better explained using threshold models, because they depend on objective evidence and thoughtful consideration of the benefit and harm of a test/treatment. Type 1 processes are unique to each decision-maker (e.g. each physician has an individual set of skills, knowledge about a condition, and knowledge gaps). Previous studies have shown that when type 1 processes dominate decisions exclusively, ordering a diagnostic test does not affect a decision; the decision is based on the automatic assessment of knowledge-based benefits and harms of the treatment

(e.g. antibiotics for a child with an ear infection) (B. Djulbegovic et al., 2018; B. Djulbegovic et al., 2012; Mukherjee, 2010). These findings explain variations in physicians' ordering of diagnostic tests and treatment patterns (due to knowledge-to-action gaps) as commonly seen in clinical practice.

DPT have been criticized due to the conceptual vagueness, lack of precision (Keren & Schul, 2009), lack of consistency between researchers (Gigerenzer, 1996), and inferential gaps (Osman, 2004). On the other hand, DPT has been praised for its simplicity and embedded heuristics when explaining binary choices and associated biases, support from empirical evidence, and its applicability to medical decisions (B. Djulbegovic et al., 2012; Mukherjee, 2010; Tsalatsanis et al., 2015).

In summary, the EUT and DPT have different properties, strengths, and limitations. Table 1 summarizes the main characteristics and differences between the EUT and DPT when applied to medical decisions. I explain below the reasons for choosing the DPT as the framework for my studies. Certainly, physicians do not need these theories to make diagnostic or therapeutic decisions. However, these theories provide a framework to explain medical decisions and ideally overcome suboptimal choices or medical errors.

The debate about different theories will continue long after this thesis, as there is no single decision model that can fit all medical contexts (B. Djulbegovic & Elqayam, 2017; B. Djulbegovic et al., 2018). I acknowledge that either EUT, prospect theory or DPT could explain therapeutic decisions under uncertainty. Experimental evidence supports the highlighted strengths of DPT, a recognized framework to explain medical

**Table 1**. Comparison between EUT and DPT in medical decisions

| Characteristics | EUT | DPT |
|---|---|---|
| Framework | Unimodal (expected utility) | Dual process (system 1 & system 2) |
| Decisions based on mathematical computations to optimize outcomes | Normative (decisions should be based on computations to maximize EU) | Descriptive (sometimes decisions are based on computations) |
| Mathematical expression | $[EU(A)=\sum o \in O PA(o)U(o)]$* | $[p_{tt} \leq p < p_{rx}]$† |
| Accounting for social aspects (e.g. accessibility, medical insurance) | No | Takes into consideration some social aspects of treatment decisions |
| Accounting for emotional aspects | Risk preferences (e.g. risk aversion) | Risk preferences, beliefs, ambiguity |
| Accounting for physician's expertise and experience | No | Yes |
| Consideration of benefits/harms | Yes | Yes |
| Consideration of patient-centered outcomes | No | Yes |
| Decision-maker choice | Should always be based on the outcome with the highest expected utility (EU) | Also based on social, emotional, experience of the decision maker |
| Strengths | - Simple to use<br>- Practical estimations | - Considers other relevant aspects of the decision-making process (see above) |
| Critiques & Limitations | - Limited consideration of emotions<br>- Does not account for patient's preferences | - Conceptual vagueness<br>- Lack of precision defining system 1 and 2<br>- Limited neural evidence |
| Rationale for its application in medicine | Facilitate decisions based on a simple clinical parameter | Overcome the limitations of the EUT |

* where O is the set of outcomes, PA($o$) is the probability of outcome o conditional on A, and U($o$) is the utility of $o$. † where $p_{tt}$ represents testing threshold and $p_{rx}$ represents the treatment threshold. The probabilities $p_{tt}$ and $p_{rx}$ are functions of a decision maker's attitudes towards treatment benefits and harms as well as harms of testing, both derived from type 1 and type 2 cognitive mechanisms.
The probability of disease, $p$ can be estimated by statistical evidence, and by the physician's intuition and experience. Further details and examples that exceed this thesis can be found elsewhere.(Tsalatsanis et al., 2015)

decisions under uncertainty (Davis, McCaffery, Mullan, & Juraskova, 2015; B. Djulbegovic et al., 2018; Fargen & Friedman, 2014; Fargen, Leslie-Mazwi, Chen, & Hirsch, 2020; Stolper et al., 2011; Tsalatsanis et al., 2015). More importantly, it provides a better framework to explain automatic and deliberate choices under uncertainty, also accounting for other relevant components influencing medical decisions: social and emotional factors, benefits/harms weight, and physician's experience.

Medical schools and post-graduate training educate physicians to quickly recognize patterns or critical aspects of different diseases.(Mamede, van Gog, et al., 2010; Prakash et al., 2019; Sibbald, de Bruin, & van Merrienboer, 2013) Physicians apply the knowledge they have acquired from previous experience, use information available at the time of the assessment, use risk score tools, educational interventions or a combination of the above to make automatic or deliberate therapeutic decisions.(Gongora-Ortega, Segovia-Bernal, Valdivia-Martinez Jde, Galaviz-DeAnda, & Prado-Aguilar, 2012; Meinema, Buwalda, van Etten-Jamaludin, Visser, & van Dijk, 2019)

**What are the brain pathways that facilitate this decision process?**

Studies in decision neuroscience have shown that individuals make goal-directed decisions by assigning value to different options, and comparing them to make a choice (Hare, Camerer, & Rangel, 2009; Rangel, Camerer, & Montague, 2008). Under the DPT, this process occurs under system 2. Participants make goal-directed decisions using values computed in the ventromedial prefrontal cortex (vmPFC) (Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008; Plassmann, O'Doherty, Shiv, & Rangel, 2008). The avoidance of automatic responses (self-control) involves the modulation

of value signals encoded in the vmPFC by the dorsolateral prefrontal cortex (DLPFC) (Carter & van Veen, 2007; Ochsner & Gross, 2005; Padoa-Schioppa & Assad, 2006). The translation of these studies into clinical practice may suggest that physician's goal-directed decisions (e.g. ordering a diagnostic, selecting a treatment option) would be based on values (e.g. reliability of a test, efficacy of a treatment) computed either in the vmPFC or other brain structures. Different brain pathways, involving DLPFC (and other brain regions) would modulate automatic medical decisions by either continuing with the same treatment (status quo) or making a more valuable treatment choice. To the best of my knowledge there are no previous studies assessing value-based options among practicing physicians, especially value-based treatment decisions.

In the present work, I elected to use DPT as the theoretical framework to explain decisions under uncertainty. As mentioned, DPT helps explain binary choices in the medical field (i.e. automatic vs. analytical decisions to treat or not to treat, test or not to test), accounts for cognitive biases and personality traits, and includes the conceptualization of treatment choices when assessing TI (i.e. treatment escalation vs. continuing the same treatment). Furthermore, DPT provides the rationale for creation and application of my educational intervention (the "*traffic light system")* (Saposnik, Mamdani, et al., 2019), the design of our studies, and the background to elucidate mechanistic pathways to overcome TI. For example, I show that increased arousal is associated with TI.

As our understanding of disease prevention (and progression) continues to be refined (particularly in multiple sclerosis), personalized medicine must adapt to account for physician's risk preferences, expertise, and other factors involved in therapeutic

decisions (Rotstein & Montalban, 2019; Saposnik, Sempere, et al., 2016a). The present work on therapeutic inertia provides the foundations for future research in this area.

## 1.2 Multiple sclerosis: The paradigm of therapeutic decisions under uncertainty

### What is Multiple sclerosis and the current treatment?

Multiple sclerosis (MS) is an autoimmune inflammatory disease of the central nervous system (brain, spinal cord) and a leading cause of disability and diminished quality of life in young individuals worldwide (Rotstein & Montalban, 2019; Vidal-Jordana & Montalban, 2017). MS attacks myelin, the protective covering of the nerves, causing inflammation and damage. Myelin is a critical component for the transmission of nerve impulses through nerve fibers (Matute-Blanch, Montalban, & Comabella, 2017; Rotstein & Montalban, 2019).

MS can cause symptoms like a lack of coordination, weakness, impaired sensation, vision problems, bladder problems, generalized fatigue, cognitive impairment and mood changes. Its effects can be physical, emotional, and financial (Burks, Marshall, & Ye, 2017; Ernstsson et al., 2016; Giovannoni, 2018; Matute-Blanch et al., 2017). The course of MS is typically progressive, characterized by recurrent neurological events (e.g. relapses) with complete or partial recovery (Matute-Blanch et al., 2017; Rotstein & Montalban, 2019).

The field of MS research has seen significant changes over the last several years. For example, new therapies (i.e. monoclonal antibodies) tested in clinical trials have shown to reduce the number of relapses, the number of brain lesions and improving

the functional status and quality of life of patients (Feinstein, Freeman, & Lo, 2015; Mahad, Trapp, & Lassmann, 2015; Rotstein & Montalban, 2019). Currently, there are over 16 approved disease-modifying therapies (DMTs) for MS with varying dosage and administration routes (oral, injectable, and infusion), with different safety and efficacy profiles. Injectable agents (e.g. interferons, glatiramer) dominated MS care for over two decades (Mark S. Freedman, Selchen, Prat, & Giacomini, 2018; Giovannoni, 2018; Rotstein & Montalban, 2019). Injectable agents for MS promise an approximately 30% reduction on the annual relapse rate and have similar risk–benefit profiles (e.g. flu-like symptoms being the most common side-effects). The introduction of oral agents (e.g. Fingolimod, Dimethyl-fumarate) and new humanised monoclonal antibody (e.g. Natalizumab, Alemtuzumab, Ocrelizumab) administered by infusions have opened another therapeutic avenue for patients and clinicians. Monoclonal antibodies are more effective treatment options (e.g. an ~80% reduction in the annual relapse rate) but carry the risks associated with modulation of the immune system (e.g. more serious side effects related to infections, risk of a fatal leukoencephalopathy, progressive damage or inflammation of the white matter of the brain at multiple locations) (D. Ontaneda, Fox, & Chataway, 2015; Sormani & Bruzzi, 2015). Further details of MS can be found https://mssociety.ca/about-ms/what-is-ms.

**The paradigm of therapeutic decisions under uncertainty**

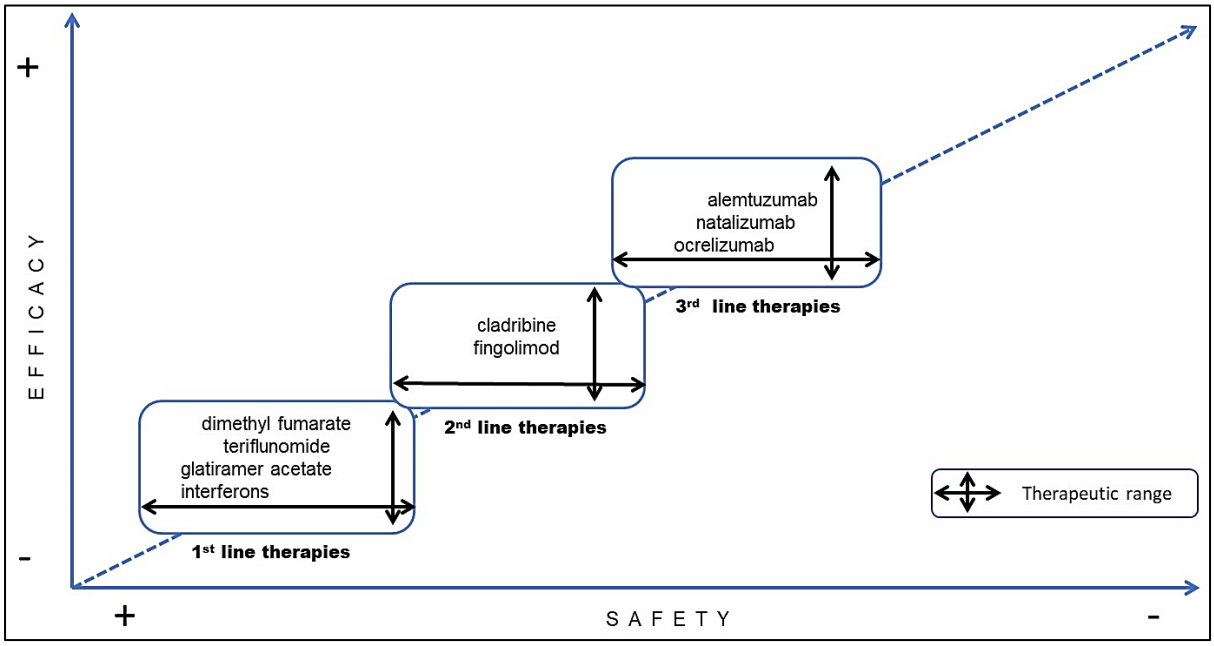
Multiple sclerosis represents an excellent model for the study of treatment decisions for chronic and progressive medical conditions under uncertainty (Oh, Vidal-Jordana, & Montalban, 2018). The wide availability of treatment options ranging from low efficacy/low risk of side effects (e.g. interferons) to high-efficacy/higher risk of serious side effects (e.g. monoclonal antibodies), represents a daily challenge that physicians

face (Mark S. Freedman et al., 2018; Giovannoni, 2018; Rotstein & Montalban, 2019). Given the several 'unknowns' (e.g. an individual patient's response and tolerance of side effects under different treatments), each treatment decision bears some degree of uncertainty. Physicians' cognitive biases, beliefs, and risk preferences may also influence their therapeutic choices (Blumenthal-Barby & Krieger, 2015; Saposnik, Redelmeier, Ruff, & Tobler, 2016).

In this current landscape, physicians strive to balance efficacy and safety with a large and diverse therapeutic arsenal. Treatments are tailored based on: i) disease activity level (clinical data), ii) individual patient characteristics/preferences, and iii) personal expertise/preference (Daniel Ontaneda, Cohn, & Fox). In this context, it is not surprising that the uncertainty related to switching to a new treatment may lead to the default choice (e.g. continuing with the 'known' treatment profile), or maintaining the status quo (e.g. therapeutic inertia) (Fleming, Thomas, & Dolan, 2010; Saposnik & Montalban, 2018; Shevchenko, von Helversen, & Scheibehenne, 2014). This scenario combining a broad spectrum of treatment options, physician's cognitive biases, management of uncertainty provides a unique opportunity to improve our understanding on how physicians make therapeutic decisions.

I have chosen to focus on therapeutic decisions in MS because of well-established safety and efficacy profiles amongst first (interferons, glatiramer), second (Fingolimod, Cladribine), and third line (monoclonal antibodies) medications in an evolving landscape (see Figure 1 below).

**Figure 1. Landscape of MS treatment**




*This figure represents the current landscape of MS treatment. Each box represents first-, second- or third-line therapies when consider treatment escalation. Different MS drugs are showed represented inside each box. Arrows illustrate the therapeutic range, meaning the balance between the efficacy and safety profile. Participants rated the level of agreement with this figure using a Likert scale with 0 being the lowest and 10 the highest scores.*

Furthermore, I have access to neurologists and MS experts across the world to conduct different studies, test my hypothesis, and answer the proposed research questions. However, multiple sclerosis is not the only progressive and chronic medical condition, and lessons learned from the study of MS may be extended across medical disciplines. Hypertension, diabetes, and high cholesterol represent other highly prevalent chronic diseases that lead to the progression of atherosclerosis (buildup of fats, cholesterol, and plaque on artery walls). Like MS, there are a broad spectrum of treatment options for hypertension, diabetes and high cholesterol (Diaz Rodriguez et al., 2014; Lebeau et al., 2014; Lebeau et al., 2016; Manski-Nankervis et al., 2017; Mata-Cases et al., 2018; Milman, Joundi, Alotaibi, & Saposnik, 2018). Lessons learned from MS can be extended and applied to other chronic medical conditions.

## 1.3 Therapeutic Inertia in MS care: Conceptualization

Despite recent therapeutic and technologic advances in medicine, we have limited information on how physicians make decisions. Medical schools do not properly train future doctors in decision making and educate them on risk-management strategies (Dijkstra, Pols, Remmelts, Rietzschel, et al., 2015; Mamede et al., 2014). As a result, physicians are vulnerable to exhibit cognitive biases that may influence diagnostic and therapeutic decisions. One of the most relevant cognitive biases in medicine is the status-quo.(Blumenthal-Barby & Krieger, 2015; Fleming et al., 2010; Saposnik, Redelmeier, et al., 2016; Thaler & Sunstein, 2008) The most concerning consequence of status-quo is therapeutic inertia (TI).(Corathers & DeSalvo, 2020; Redelmeier & Shafir, 1995; Ritov & Baron, 1992)

Therapeutic inertia (TI) emerged as a concept to classify suboptimal decisions when treatment goals are unmet in three prevalent general medical conditions: hypertension, diabetes and chronic obstructive pulmonary disease (Cooke, Sidel, Belletti, & Fuhlbrigge, 2012; Khunti et al., 2016; Mata-Cases et al., 2018; Milman et al., 2018; Ogura & Harada-Shiba, 2016). The criteria to define ‘unmet therapeutic goals or targets’ are based on best practice recommendation guidelines for specific medical conditions. TI is being referred as the tendency of physicians to continue the same treatment (i.e. status-quo) despite clinical evidence of disease progression.(Mohan & Phillips, 2011; Reach, 2014; Saposnik, Sempere, et al., 2016a) The invoked explanation for TI, status-quo and other cognitive biases resides in physicians’ limited training in risk management and formal learning in medical decision-making.

TI may be the consequence of physicians' limited training in risk management and formal learning in medical decision-making. TI is a novel concept in MS care. Before 2016, there were no previous studies evaluating therapeutic inertia and its associated factors in MS care (Saposnik, Sempere, et al., 2016a). Given the gaps in the conceptualization of therapeutic inertia, I have chosen to focus on TI in MS for this work. Previous studies have shown that insufficient knowledge integration and knowledge-to-action gaps are among the most common explanations for suboptimal therapeutic decisions, medical errors or TI (Fleming et al., 2010; Ioannidis & Lau, 2001; Maier, Ernst, & Steinhauser, 2019; Makary & Daniel, 2016; O'Connor, Sperl-Hillen, Johnson, Rush, & Biltz, 2005; Stiegler & Ruskin, 2012).

### 1.3.1 What are the critical elements to define TI in MS care?

National and international best practice guidelines have shown that the critical elements used to define TI include the *clinical course* (i.e. the presence of new relapses despite being on DMTs), *neuroimaging* (i.e. presence of new or active lesions on a patient's MRI while being on DMTs) and the *functional status* of patients (i.e. defined by the validated Expanded Disability Status Scale [EDSS]).(Hobart et al., 2019; Rotstein & Montalban, 2019) This concept is also known in the MS literature as 'no evidence of disease activity' (NEDA) to illustrate patient's stability and guide treatment (i.e., the combination of lack of a clinical relapse, stable functional status and stable MRI suggests an effective treatment response) (Giovannoni, 2018; Giovannoni et al., 2015). Conversely, there is consensus that the presence of clinical relapse within a time-period and new lesions on MRI warrant treatment escalation. (Giovannoni, 2018; Giovannoni et al., 2015; Montalban et al., 2018; Rae-Grant et al., 2018)

Although physicians may also consider other factors when deciding to escalate treatment, I lead a recent international study comprising over 300 MS neurologists practicing in 25 countries. The study applied a conjoined design and confirmed the hierarchy of these three components. Other factors (e.g., age, gender, severity or localization of symptoms, pregnancy status) has a much lower (<10%) relative importance in therapeutic decisions (Saposnik et al. Manuscript in preparation). As shown in Figure 1, the landscape of DMTs for the treatment of MS includes first-line therapies (beta interferons, glatiramer acetate, teriflunomide and dimethyl fumarate) and second-line therapies (fingolimod) and third-line therapies (monoclonal antibodies such as natalizumab, alemtuzumab, ocrelizumab).

The proposed framework of therapeutic decisions in MS care does not consider the concept of shared decision-making (SDM), defined as the process by which doctors and patients work together to make choices after discussing benefits and risks of appropriate treatment options (Glyn Elwyn et al., 2012). Of note, shared decisions carry some challenges. For example, a critical aspect of SDM can only take place once physicians identify the best course of action before offering and discussing treatment options with patients (Hoffmann, Jansen, & Glasziou, 2018; Légaré & Witteman, 2013). The simulated case-scenarios were designed to facilitate binary decisions that match clinical practice (therapeutic choices aligned with treatment escalation vs. alternatives under the same line of therapy with no treatment escalation). As a result, the case-scenarios presented to assess TI should leave treatment escalation as the only correct choice for participants.

In summary, the most critical elements that guide physicians towards treatment escalation (and thereby towards avoiding TI) include the integration of three

components: i) the clinical course, ii) results from neuroimaging, and iii) the functional status of patients. The designed simulated-case scenarios to investigate the concept of TI were based on internationally accepted and established criteria for treatment escalation.

### 1.3.2 What are the consequences of TI in MS care?

Previous studies in MS care revealed that a proactive management approach (e.g. including earlier use of high-efficacy DMTs and close monitoring of the clinical and radiological response to treatment) is associated with slow disease progression, lower disability, lesser cognitive impairment, and diminished MRI activity (Duquette, Giacomini, Bhan, Hohol, & Schecter, 2016; Noyes & Weinstock-Guttman, 2013; Prosperini et al., 2012; Sormani et al., 2013) and lower health care costs (M. S. Freedman et al., 2019; Gani et al., 2008). Meta-analysis confirmed that second- and third-line DMT are the best available choices for preventing disease progression in patients with MS (M. S. Freedman et al., 2019; Tramacere, Del Giovane, Salanti, D'Amico, & Filippini, 2015).

In summary, the consequences of TI include increased likelihood of disease progression, worsened functional status, loss of independency for activities of daily living, higher probability of cognitive impairment with associated increased health care costs (M. S. Freedman et al., 2019; Gani et al., 2008; Jongen et al., 2015; McCrone, Heslin, Knapp, Bull, & Thompson, 2008; Ness, Haase, et al., 2020; Ness, Schriefer, et al., 2020; Richards, Sampson, Beard, & Tappenden, 2002).

### 1.3.3 Operational definition of TI in the management of Multiple Sclerosis

The operational definition of TI in MS includes the combination of evidence of *clinical relapses* and *disease activity* on the MRI of the brain/spine despite receiving a DMT. Scientific organizations and regulatory drug agencies (Federal Drug Agency in the United States of America [FDA], European Medicines Agency [MA]) recommend treatment escalation from interferons (first-line therapy) to fingolimod or monoclonal antibodies in patients who have had at least one clinical relapse in the previous year and either 5 or more new lesions or ≥1 gadolinium-enhancing T1 lesion on brain MRI (Bermel et al., 2013; Cristiano et al., 2018; Mark S. Freedman et al., 2018; Garcia Merino et al., 2017; Prosperini et al., 2014; Rae-Grant et al., 2018; Sormani et al., 2013). These recommendations are consistent with evidence regarding the risk of treatment failure among patients receiving interferon-β (Mark S. Freedman et al., 2018; Rotstein & Montalban, 2019; Sormani et al., 2016).

### 1.3.4 Other concepts and operational definitions relevant to my thesis

Cognitive psychologists suggests that stress, anxiety, and uncertainty trigger the expression of biases and personality traits that influence the decision-making process (Paulus & Yu, 2012; Platt & Huettel, 2008; L. Zhang, Wang, Zhu, Yu, & Chen, 2015). As mentioned, most medical decisions involve uncertainty commonly associated with having imperfect information (e.g. diverse clinical presentation for the same condition, unknown co-morbid factors, unknown treatment response for an individual, unknown medication adherence or tolerance to side effects). In the present work I applied the concept of decisions under uncertainty as commonly used by psychologists and economists. I assessed several concepts that may influence medical decisions under uncertainty, including risk preferences, ambiguity aversion, and physician's reaction

to uncertainty (M. Gerrity, White, DeVellis, & Dittus, 1995; Levy, Snell, Nelson, Rustichini, & Glimcher, 2010; Platt & Huettel, 2008).

Uncertainty comprises two different terms: risk and ambiguity. Ambiguity is a term used when the probability of an event or outcome is unknown (Ellsberg, 1961; Platt & Huettel, 2008). Aversion to ambiguity is defined as dislike for events with unknown probability over events with known probability (Levy et al., 2010). For example, an ambiguity-averse individual would rather choose a treatment where the probability of benefits or side effects are known (even if these are somewhat unfavourable) over one where these probabilities are unknown.

In contrast, risk involves individual choices when the probability of an outcome is known. Risk aversion is defined as the tendency to prefer safe payoffs over probabilistic payoffs when the expected value of both options is identical (Camerer & Weber, 1992; Levy et al., 2010). A risk-averse patient would thus prefer a treatment that provides a small known improvement with certainty over a treatment that provides equal chance of a large improvement or no improvement (50/50). I evaluated risk aversion by identifying the safe amount for which a participant was indifferent between the safe and the risky option (Christopoulos, Tobler, Bossaerts, Dolan, & Schultz, 2009). This level is called the certainty equivalent, and reflects the value associated with the risky option to facilitate comparison between participants.

I also used two standardized surveys to assess physicians' willingness to take risks and physician's tolerance to uncertainty. The German Socio-Economic Panel (SOEP) is a comprehensive multi-dimensional survey (and database) that collects information to better understand human behavior and decision making in varying social and institutional settings and can be compared longitudinally. Although the SOEP survey has not been tested as a determinant of medical decisions, previous studies have

shown an association between personality traits and social outcomes (Bagnjuk, Konig, & Hajek, 2019; Gohlmann, Schmidt, & Tauchmann, 2010; Hajek, Bock, & Konig, 2017). As a result, I selected a component from the validated SOEP that specifically evaluates participant's willingness to take risks in different domains (financial matters, own health, car driving, own occupation, sports and leisure activities) (Dohmen et al., 2011). The second survey measured physicians' tolerance to uncertainty in patient care, using the physician's reaction to uncertainty test (M. S. Gerrity, DeVellis, & Earp, 1990). A shorter version following a factor analysis comprises five questions showing reliable psychometric properties (α-Cronbach 0.90). Participants rate the level of agreement with the following statements from 0 (strongly disagree) to 5 (strongly agree): i) the uncertainty of patient care often troubles me; ii) I find the uncertainty involved in patient care disconcerting; iii) I usually feel anxious when I am not sure of the diagnosis; iv) uncertainty in patient care makes me uneasy; and v) I am quite comfortable with the uncertainty in patient care. Note that the last item is reverse coded for consistency. After participants provided a rating for each question, all are added to obtain a total score (Cunningham, Bonham, Sellers, Yeh, & Cooper, 2014). Previous studies have shown that physician's low tolerance to uncertainty was associated with higher resource utilization and patients being recalled for studies (Allison et al., 1998; Carney et al., 2007). Further details of the operational definitions for each of these measures are summarized in section 2.2 (Saposnik, Sempere, et al., 2017).

**What is the rationale for using pupil measures in my studies for this thesis?**

Pupil dilation is a marker of autonomic arousal (C.-A. Wang et al., 2018). Tonic and phasic pupil responses are modulated by a constant balance between

parasympathetic and sympathetic pathways. From the behavioral perspective, arousal states are mediated by the locus coeruleus. fMRI studies showed a correlation between pupil size and the activation of the locus coeruleus (Clewett, Huang, Velasco, Lee, & Mather, 2018; Peter R Murphy, O'Connell, O'Sullivan, Robertson, & Balsters, 2014). Importantly, autonomic arousal is an established proxy measure of uncertainty (Geng, Blumenfeld, Tyson, & Minzenberg, 2015; Lavin, San Martin, & Rosales Jubal, 2013; Urai, Braun, & Donner, 2017). Altogether, pupil dilation is a promising indicator of how physicians manage uncertainty when making therapeutic decisions. Furthermore, pupil dilation could also measure how the TLS educational intervention modulates changes in physician's decisions under uncertainty.

## 1.4 Five Outstanding questions

I was interested in addressing the following questions:

**Question 1: What are the most common cognitive biases and how do they affect physicians' decision-making? More specifically, have there been studies on status-quo (also called default bias) in therapeutic decisions?**

The health sector shares commonalities with industrial sectors including vulnerability to cognitive biases and human errors.(Stripe, Best, Cole-Harding, Fifield, & Talebdoost, 2006; Zeltser & Nash, 2010) Several cognitive factors and biases affect physicians' thinking process. However, it is not known the prevalence of cognitive biases and their influence in treatment decisions.

**Question 2: What is the prevalence and what are the most common factors associated with TI amongst neurologists with expertise in MS care?**

Therapeutic inertia (TI) is a common phenomenon in the management of chronic medical conditions: hypertension, diabetes and chronic obstructive pulmonary disease (Cooke et al., 2012; Khunti et al., 2016; Mata-Cases et al., 2018; Milman et al., 2018; Ogura & Harada-Shiba, 2016). The most common factors associated with TI were related to physician's training, limited education or expertise in these three medical conditions leading to inadequate treatment. I introduced the concept of TI in MS care at the beginning of my thesis in 2016.(Saposnik, Sempere, et al., 2016a) The consequences of TI include poor patient outcomes, diminished quality of life, and loss of productivity.(Gupta, Goren, Phillips, Dangond, & Stewart, 2014; Rotstein & Montalban, 2019). My goal was to determine the factors associated with TI in neurologists with expertise in MS care.

**Question 3: What is the consequence of a physician's tolerance to uncertainty, ambiguity aversion or therapeutic inertia?**

Previous studies have shown that a physician's low tolerance to uncertainty may lead to medical errors (M. S. Gerrity et al., 1990; Yee, Liu, & Grobman, 2014). Status quo has also been associated with decisions under uncertainty and medical errors (Aberegg, Haponik, & Terry, 2005; Fleming et al., 2010). However, there have been no previous studies on the influence of physicians' tolerance to uncertainty, aversion to ambiguity or status-quo in MS care.

**Question 4: What is the efficacy of the traffic light system (TLS) approach to ameliorate TI in MS care?**

Educational interventions have been designed to optimize knowledge integration and bridge knowledge-to-action gaps for complex medical decisions (e.g. diagnostic challenges, varying risk categories, availability of multiple agents with a broad range of safety/efficacy ratios) (Dijkstra, Pols, Remmelts, & Brand, 2015). One such promising tool was the Traffic Light System (TLS), which links a warning sign (e.g. red color) to a prognosis (e.g. risk of disease progression) guiding physicians to make a decision (e.g. treatment switch, hospital admission, etc.) (M. S. Murphy & Baker, 2014; X. Zhang, Liu, Gu, Wang, & Chen, 2020).

The TLS relies on a well-established and cross-cultural learned link between a color (e.g. red) and an action (e.g. stop, think, make a decision) to modify the natural chain of expected events (M. S. Murphy & Baker, 2014; Sonnenberg et al., 2013). It uses existing brain pathways such as the inferior frontal gyrus/dorsolateral prefrontal cortex (regions implicated in self-control), the posterior cingulate cortex and the ventromedial prefrontal cortex implicated in suppressing automatic responses and updating the valuation system.(L. Enax, Hu, Trautner, & Weber, 2015; L. Enax, Krapp, Piehl, & Weber, 2015). Using neurocognitive methods, Enax et al. showed that the TLS facilitates healthier food choices by interrupting automatic behavior and triggering a re-evaluation processes (Laura Enax et al., 2016). An fMRI study showed that TLS labels enhance the coupling between brain regions associated with valuation (i.e. ventro-medial prefrontal cortex) and self-control (L. Enax, Krapp, et al., 2015).

Evidence from the literature suggests that medical decisions leading to TI are likely related to knowledge-to-action gaps (Gongora-Ortega et al., 2012). The design and application of a TLS in MS care would provide a unique opportunity to overcome

knowledge-to-action gaps by taking the advantage of existing brain pathways that suppress automatic responses and update the valuation system (L. Enax, Hu, et al., 2015; A. C. Phillips et al., 2016). I therefore designed and tested the feasibility and efficacy of a simple educational intervention (supported by a theoretical neuroeconomic framework- TLS) in a randomized controlled trial to overcome TI in MS care.

**Question 5: What is the interaction between an arousal response on the effects of an educational intervention (e.g. application of the traffic light system approach) to decrease TI in MS?**

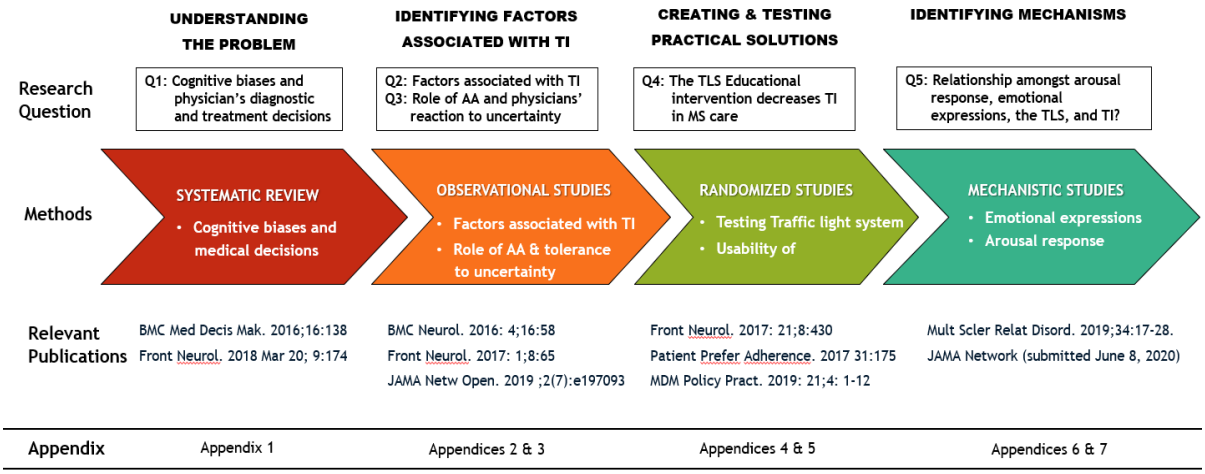
My final question aimed to evaluate how the TLS intervention reduce TI and link this relationship with the autonomic arousal response via the aforementioned brain pathway. Recent studies showed pupil dilation is a marker of central arousal, suggesting that individual arousal responsivity may affect decisions under uncertainty (Mathot, 2018; Urai et al., 2017). This hypothesis is supported by reports of an association between the level of arousal (expressed as phasic pupil response) and suboptimal or incorrect decisions in non-therapeutic settings (de Gee, Knapen, & Donner, 2014; P. R. Murphy, Vandekerckhove, & Nieuwenhuis, 2014; Urai et al., 2017). The decision-making process is mediated by the integration of different pathways involving the noradrenergic locus coeruleus (a core component of the brain's arousal system), which is directly connected with self-control, automatic responses and valuation systems (Cohen, 2005; Ekman & Friesen, 2003; Joshi, Li, Kalwani, & Gold, 2016; Mathot, 2018; Phelps, Lempert, & Sokol-Hessner, 2014; C. A. Wang & Munoz, 2015). However, it is entirely unclear if and how individual differences in arousal responsivity impact treatment decisions in the real world and in real life

scenarios. For example, what is the arousal state when expert neurologists are exposed to live therapeutic decisions? I therefore designed an experiment to assess differences in arousal responses among expert neurologists while they made therapeutic decisions based on auditorily presented simulated case-scenarios. I hypothesized that physician's exposure to uncertainty (elicited by the exposure to simulated medical scenarios) would generate an anticipatory arousal response that could lead to therapeutic inertia. I hypothesized that the TLS educational intervention would ameliorate TI via a reduction in the arousal state of participants. Participant's expertise, years of practice, and possibly tolerance/aversion to uncertainty and/or ambiguity are hypothesized to be cofactors that could contribute to the reduction of TI in MS care.

# 2. Summary of the Experimental Approach

To answer these five questions, this thesis draws on four methodological tools: 1) a systematic review of the literature regarding the association of physician's cognitive biases and medical decisions; 2) a series of prospective cohort studies to assess most common factors associated with TI in expert neurologists; 3) a randomized trial to test the efficacy on my behavioral-based educational intervention applying the TLS; and 4) application of eye-tracking and facial recognition systems to investigate the relationship between emotional expressions and TI and arousal state with the TLS and TI. I break down these five questions into 6 research studies (Figure 2).

**Figure 2. Integration of research questions, studies, and publications**




MS: Multiple sclerosis; AA: ambiguity aversion; TI: Therapeutic inertia; TLS: Traffic light system. *This figure illustrates the integration of the rationale for each study, research questions, methodology, and publications related to the present work.*

Study 1 focuses on a comprehensive systematic review assessing physician's cognitive biases that may affect medical decisions (Saposnik, Redelmeier, et al., 2016). Specifically, I have four objectives: 1) identify the most common cognitive factors and biases affecting physicians in medical encounters or simulated case-scenarios, 2) evaluate the influence of cognitive biases on diagnostic accuracy, or

therapeutic or management errors, 3) determine their impact on patient outcomes, and 4) identify literature gaps which could lead to recommendations that advance our understanding of therapeutic inertia (see Appendix 1).

To answer research Questions 2 and 3, I conducted two prospective cohort studies (Figure 2). Initially, I assessed factors associated with TI in a cohort of neurologists from Spain. I also experimentally assessed the influence of ambiguity aversion and physician's reaction to uncertainty on TI (Saposnik, Sempere, et al., 2017). I then tested factors associated with TI in other countries (Argentina, Chile, Canada), applying the same study design to verify reproducibility and determine the generalizability of my findings (Almusalam et al., 2019). See Appendices 2 and 3.

My fourth research question focused on the development and pilot testing of our TLS-based educational intervention (Question # 4, Appendix 4)(Saposnik, Maurino, et al., 2017). I next assessed the efficacy of the TLS intervention in a larger randomized controlled trial (Appendix 5).

I also tested the acceptance of the TLS intervention by measuring the validated usability score (Lewis & Sauro, 2009; Saposnik, Tobler, et al., 2018).

My fifth study focused on evaluating the relationship between arousal states (measured by pupil dilation), application of the TLS, and therapeutic decisions. Specifically, I was interested in evaluating how the TLS educational intervention modified the way physicians manage decisions under uncertainty. As mentioned, autonomic arousal (measured by pupil dilation from baseline) is an established proxy measure of uncertainty.(Geng et al., 2015; Lavin et al., 2013; Urai et al., 2017) To test this relationship, neurologists with expertise in multiple sclerosis used the educational

intervention during their evaluation of simulated case-scenarios. I aimed to: i) evaluate the relationship between arousal responses and TI, ii) investigate how the validated TLS educational intervention (Saposnik, Mamdani, et al., 2019) affected arousal responses and TI, and iii) assess whether arousal responses modulated the association between the educational intervention and TI (i.e. mediation analysis) (see Appendix 6).

Given that emotions influence decision-making, I also tested the relationship between emotions and affective states (as captured by muscle facial activity and emotional expressions) and TI amongst neurologists caring for MS patients when making therapeutic decisions. I used a validated machine learning algorithm from AFFDEX software to code for facial muscle activations, and predefined mapping to emotional expressions (disgust, fear, surprise) (Saposnik, Oh, et al., 2019) (see appendix 7).

Owing to this combination of methods and the sequential design of my studies, I was able to investigate these five questions systematically, beginning with the assessment of literature gaps followed by the creation and evaluation of the TLS-based educational intervention, and concluding by evaluating how physicians handle uncertainty and how the TLS intervention change physician's decisions by measuring arousal responses.

To bring this PhD to completion has required significant persistence through adversity. This program was costly, due to the significant number of expert participants (e.g. neurologists with expertise in the management of MS) required to test the consistency of our results. The total operating funding obtained for completing these studies was CHF 232,000 (USD 242,000), which is beyond the equipment and funding available at many institutes in the world. As a practicing staff neurologist at St. Michael's Hospital, affiliated to the University of Toronto in Canada, I took a sabbatical

year to complete all required credits at UZH. Then, I returned to Canada to complete the studies related to my PhD while provide clinical service with educational, clinical and research commitments. I had multiple financial and time constraints because as a self-funded international PhD student I had to obtained financial support for each of the described studies. Furthermore, I carried out the last study (see Appendix 6) by myself, including setting-up a mobile eye-tracking lab to assess arousal states and TI. Given the limited availability of colleague neurologists to participate on my research projects, I recruited expert neurologists from across Canada, which required substantial organization and travelling. Finally, there was a single (and costly) software provider (iMotions.com) that integrates data on arousal and emotional states with simulated case-scenarios and the designed experiments (e.g.: ambiguity aversion, risk preferences of participants), as the timing and precision of synchronized time points is critical for the interpretation of results. Despite these challenges and constraints, the combination of methods and strategic design of the studies have provided us a unique insight with practical implications for neuroscientists, clinicians, policymakers, and patients with MS. By identifying the determinants of TI, testing the efficacy of the TLS- educational intervention to ameliorate the effects of TI, and describing the underlying arousal and emotional states associated with TI, I hope to improve our current understanding of medical decision-making, and offer practical solutions to a common problem that physicians face when making treatment decisions in daily practice.

## 2.1 Study 1: Cognitive Biases associated with Medical Decisions: A Systematic Review

**Methods:** In this primarily narrative systematic review (Saposnik, Redelmeier, et al., 2016) [Appendix 1], I evaluated existing evidence on the relation between cognitive biases affecting physicians and medical decisions. Under the concept of cognitive biases (e.g. framing effect, status-quo), I also included personality traits (e.g. aversion to risk or ambiguity) that may systematically affect physicians' judgments or decisions, independent of whether or not they result in immediate medical errors. We searched MEDLINE and the Cochrane Library databases for relevant articles on cognitive biases from 1980 to May 2015. I included studies conducted on physicians that evaluated at least one cognitive factor using simulated case scenarios and reported an associated outcome written in English. Data quality was assessed by the Newcastle-Ottawa scale.

Over 32 types of cognitive biases have been described (Croskerry, 2003). Importantly, some of these may reflect personality traits that could result in choice tendencies that are factually wrong, whereas others reflect decisions that are potentially suboptimal, although there is no objectively "correct" decision (e.g. risk aversion, tolerance to ambiguity). Both factors were included here. This review had four objectives: 1) to identify the most common cognitive biases by subjecting physicians to real world situations or case-vignettes, 2) to evaluate the influence of cognitive biases on diagnostic accuracy and medical errors in management or treatment, 3) to determine which cognitive biases have the greatest impact on patient outcomes, and 4) to identify literature gaps in this specific area to guide future research. After addressing these objectives, I conclude by highlighting the practical implications of our findings and by outlining an action plan to advance the field.

**Results and discussion:** Among 114 publications, 20 studies comprising 6810 physicians met the inclusion criteria. Nineteen cognitive biases were identified.

All studies found at least one cognitive bias or personality trait to affect physicians. Overconfidence, lower tolerance to risk, the anchoring effect, and information and availability biases were associated with diagnostic inaccuracies in 36.5% to 77% of case-scenarios. Five out of seven (71.4%) studies showed an association between cognitive biases and therapeutic or management errors. Of 2 (10%) studies evaluating the impact of cognitive biases or personality traits on patient outcomes, only one showed that higher tolerance to ambiguity was associated with increased medical complications (9.7% vs 6.5%; p =.004). Most studies (60%) targeted cognitive biases in diagnostic tasks, fewer focused on treatment or management (35%) and on prognosis (10%). Literature gaps include potentially relevant biases (e.g. status-quo, aggregate bias, feedback sanction, hindsight bias) not investigated in the included studies. Moreover, only 5 (25%) studies used clinical guidelines as the framework to determine diagnostic or treatment errors. Most studies (n=12, 60%) were classified as low quality.

More importantly, only 35% of studies provided information on the association between cognitive biases or personality traits and medical errors (Baldwin et al., 2005; Perneger & Agoritsas, 2011; Redelmeier & Shafir, 1995; Reyna & Lloyd, 2006; Sorum et al., 2003; Stiegler & Ruskin, 2012; Yee et al., 2014), with scarce information on their impact on patient outcomes, preventing us from making definite conclusions.(Baldwin et al., 2005; Yee et al., 2014)

Our study (Saposnik, Redelmeier, et al., 2016) added relevant information regarding the influence of cognitive biases particularly in physicians on diagnostic inaccuracies, suboptimal management and therapeutic errors, and patient outcomes. Our first

objective allowed the identification of additional biases (e.g. framing effect, decoy effect, default bias) and physician's personality traits (e.g. low tolerance to uncertainty, aversion to ambiguity), by including fourteen further studies not previously included in a previous systematic review on the same topic (Blumenthal-Barby & Krieger, 2015). I also completed a systematic quality assessment of each study using a standardized tool and identified gaps related to the influence of cognitive biases on medical errors (Wells G et al., 2013).

In summary, I highlighted the importance of recognizing physicians' personality traits and cognitive biases. I acknowledged the substantial literature gaps, limiting our understanding of the impact of cognitive biases (especially on the status-quo) and physician's personality traits on medical decisions. Although cognitive biases may affect a wide range of physicians (and influence diagnostic accuracy, management, and therapeutic decisions), their true prevalence remains unknown. I proposed the inclusion of more comprehensive study designs and the specific assessment of status-quo, ambiguity aversion, and physician's tolerance to uncertainty to evaluate their effects on medical decisions (e.g.: suboptimal decisions, therapeutic inertia) and patient outcomes in live simulated case-scenarios. I anticipate that this information would provide new insights that may affect patient outcomes (e.g. avoidable hospitalizations, complications related to a procedure or medication, request of unnecessary tests) by helping practicing physicians attenuate the prevalence of suboptimal decisions and medical errors (Andel et al., 2012; Graber, 2013; Stangierski et al., 2012).

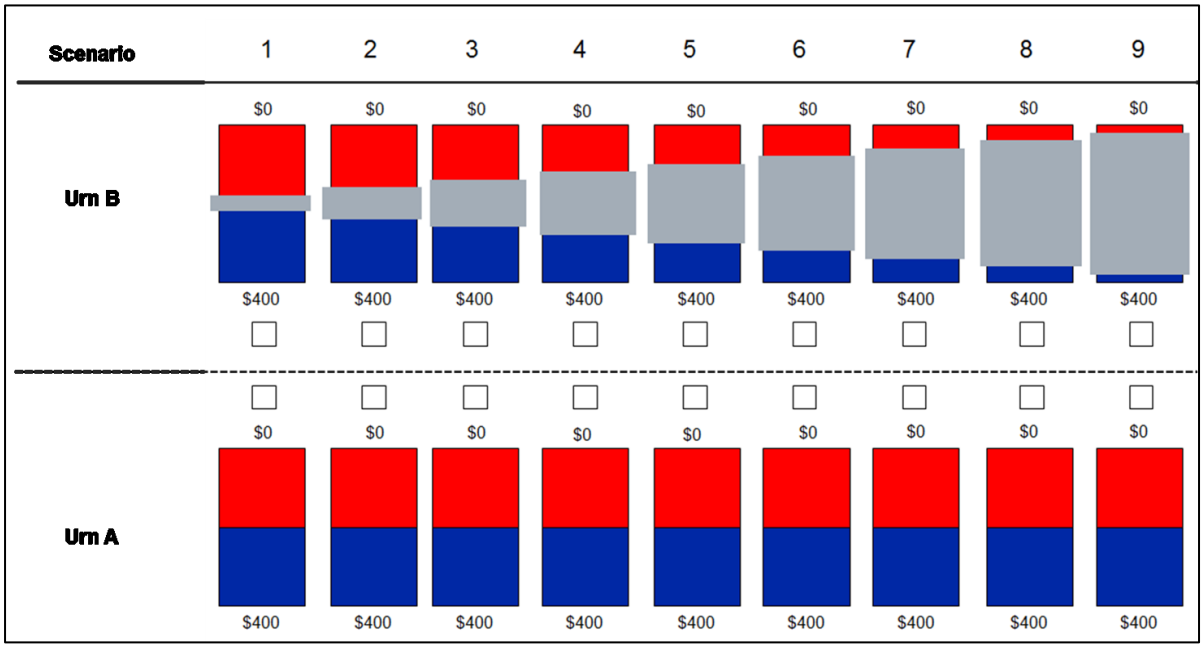
## 2.2 Study 2: Factors associated with Therapeutic Inertia

**Methods:**

I first conducted a study among practicing neurologists actively involved in the care of patients with MS from across Spain (Saposnik, Sempere, et al., 2017) [Appendix 2]. I then replicated this study in Argentina, Chile, Canada and also included a smaller sample from Spain (Almusalam et al., 2019) [Appendix 3]. Overall, the studies comprised simulated case-scenarios, 3 standardized surveys, and 4 behavioral experiments that were designed during my PhD. The simulated MS case-scenarios were derived from the most common clinical encounters as identified by experts in the field. Behavioral experiments were designed to assess risk and ambiguity aversion in the health and financial domains (Anderson & Mellor, 2008; Levy et al., 2010; Saposnik, Sempere, et al., 2017). As described, ambiguity aversion is defined as dislike for events with unknown probability over events with known probability (Levy et al., 2010). Specifically, participants were asked to choose between a visual option with known 50/50 probability of winning 400 or 0 Euros versus an option with unknown probability of the same outcomes. Grey bars represented the degree to which the winning probability was unknown (see Figure 3 below). The degree of ambiguity aversion was defined as the proportion of times participants chose the 50/50 option over the ambiguous option involving the same outcomes. As the overall level of ambiguity aversion was pronounced in this sample (mean 61.7% preference for 50/50 option, i.e. the option with known probabilities) and to avoid using an arbitrary criterion, I classified participants as highly ambiguity averse if they chose the 50/50 (known probability) option in each of the nine scenarios (Binmore, Stewart, & Voorhoeve, 2012). In order to evaluate the consistency of the relationship with the primary outcome, I also analyzed another definition of ambiguity aversion (choice of the known

probability when facing option 5, instead of the option with 50% unknown probability) (Figure 3).

**Figure 3. Design of experiments to assess ambiguity aversion**




*Participants were told to imagine two different types of urns. For urn type A, they knew that 50% of the balls were red and the other 50% were blue. For urn type B, they did not know the exact proportion of blue to red balls, with the grey bar representing the unknown proportion of balls ranging from 10 to 90% to represent different degrees of uncertainty across scenarios.*

I also assessed risk aversion as another factor that may influence clinical decisions (Gross et al., 2003). Risk aversion is defined as the tendency to prefer safe payoffs over probabilistic payoffs when the expected value of both options is identical (Camerer & Weber, 1992; Levy et al., 2010). A risk-averse patient would thus prefer a treatment that provides a small improvement with certainty over a treatment that provides a larger or no improvement with equal chance (50/50). I evaluated risk aversion by identifying the safe amount for which a participant was indifferent between the safe and the risky option (Christopoulos et al., 2009). This indifference amount, called certainty equivalent, reflects the value associated with the risky option and facilitates comparison between participants. For example, participants were asked

what would be the minimal certain payoff that they would prefer over the equiprobable gamble of winning 400 or 0 Euros (expected value of 200 euros). The degree of risk aversion of each individual corresponded to the difference of the expected value of the risky option (200 euros) minus the participant's response (proxy of certainty equivalent). A similar visual design and methodology was used to elicit aversion to risk and ambiguity in the health domains (Saposnik, Sempere, et al., 2016b). Participants were asked to choose between Treatment A (50% probability of survival) or "Treatment B" (the probability of survival is unknown) with the grey bars quantifying how much is unknown about the probability of survival.

I also used two standardized surveys to assess physicians' willingness to take risks and tolerance to uncertainty. The German Socio-Economic Panel (SOEP) is a validated survey that evaluates willingness to take risks in different domains (financial matters, own health, driving, own occupation, etc.) (Dohmen et al., 2011). I used questions of the form: "How would you rate your willingness to take risks in the following areas…."? Areas included financial matters, driving, occupation, etc. and responses could range from 0 (not at all) to 10 (very much).

The second survey measured physicians' tolerance to uncertainty in patient care, using the reaction to uncertainty test (M. S. Gerrity et al., 1990). The short version comprises five questions to be rated from 0 to 5 that when added gives a total score (Cunningham et al., 2014; M. Gerrity et al., 1995). Tolerance to uncertainty was analyzed as a continuous variable and categorical by the median split of the total score.

The primary outcome of the studies was prevalence of therapeutic inertia represented as the proportion of participants with TI in at least one simulated case-scenario. I also created the TI score, which was calculated by dividing the number of case scenarios

where participants showed TI over the number of case scenarios that were designed to measure TI (n=8). The higher the score the higher the degree of TI (Saposnik, Montalban, et al., 2018; Saposnik, Sempere, et al., 2017). Secondary outcome measures included the association between tolerance to uncertainty, risk aversion, and the SOEP surveys on the one hand with TI and therapeutic decisions on the other hand.

**Results and discussion:**

In the initial study from across Spain (Saposnik, Sempere, et al., 2017), TI was present in 68.8% (66/96) of participants. MS expertise and a higher number of MS patients seen per week were associated with a significantly lower risk of TI ($p<0.01$). Linear regression analysis suggests that the assessment of 10 more MS patients per week (from a baseline of 16) was associated with lower risk of TI (adjusted coefficient -10.2; 95%CI -18.4 to -2.0). Ambiguity aversion in the financial domain was associated with TI (86.4% vs 63.5%; $p=0.042$). Given the high correlation between specialty status with the number of MS patients seen per week (spearman 0.52; $p<0.0001$), we kept the former in the multivariable models.

Multivariable logistic regression analysis showed that high aversion to ambiguity in the financial domain was the strongest predictor of TI (adjusted OR 7.39; 1.40-38.9). Ambiguity aversion in the health domain was not associated with TI (76.9% vs. 65.7%; adjusted OR 1.79, 95%CI 0.61-5.25).

Low tolerance to uncertainty (based on the validated physician's reaction to uncertainty survey) was associated with higher prevalence of TI (85.4% vs. 56.4%; adjusted OR 4.73, 1.63-13.7). The association between TI and low tolerance to uncertainty was independent of the association between TI and high ambiguity

aversion. Conversely, willingness to take risk in multiple domains (as measured by the SOEP survey) were not associated with TI.

In the comparative study across four countries (Almusalam et al., 2019), 226 neurologists with expertise in MS care agreed to participate. The completion rate was 86.3% (195/226). The prevalence of TI was 72.8% (142/195); similar to the previous study in Spain (Saposnik, Sempere, et al., 2017). The mean TI score for the accountable 8 case-scenarios was 1.68 (±1.5), suggesting that for every 10 case-scenarios with moderate to high risk of disease progression, there would be two suboptimal decisions (e.g. lack of treatment intensification for a more effective drug) when warranted by best practice guidelines. The TI score (SD) in the Canadian group was significantly lower compared to groups from other countries (0.98 ± 1.15 vs. 1.95 ± 1.52; $p<0.001$).

Other than participants' country of origin, the multivariable analysis revealed that higher number of MS patients per week (OR 0.44; 95%CI 0.22-0.88) and years of practice (OR 0.93, 95%CI 0.86-0.99) were associated with lower likelihood of TI, whereas aversion to ambiguity was associated with two-fold higher likelihood of TI (OR 2.25; 95%CI 1.02-5.0) (Almusalam et al., 2019).

In conclusion, these international studies showed that 7 out of 10 neurologists with expertise in MS care make suboptimal decisions in at least one out of five simulated encounters. The most common factors associated with higher risk of TI included: lower expertise in MS care (e.g. lower years of experience and lower number of MS patients seen per week), country of practice, and higher aversion to ambiguity.

## 2.3 Study 3: Overcoming therapeutic inertia

This third group of studies assessed the feasibility and efficacy of my educational intervention applying the TLS. I proposed a specific strategy to facilitate physician's identification of patients with disease progression; those who require treatment intensification. In the first pilot study, I evaluated the feasibility of the TLS educational intervention (Saposnik, Maurino, et al., 2017) [Appendix 4]. In the second, larger study I evaluated the efficacy of our educational intervention in reducing TI (Saposnik, Mamdani, et al., 2019) [Appendix 5].

**Methods:**

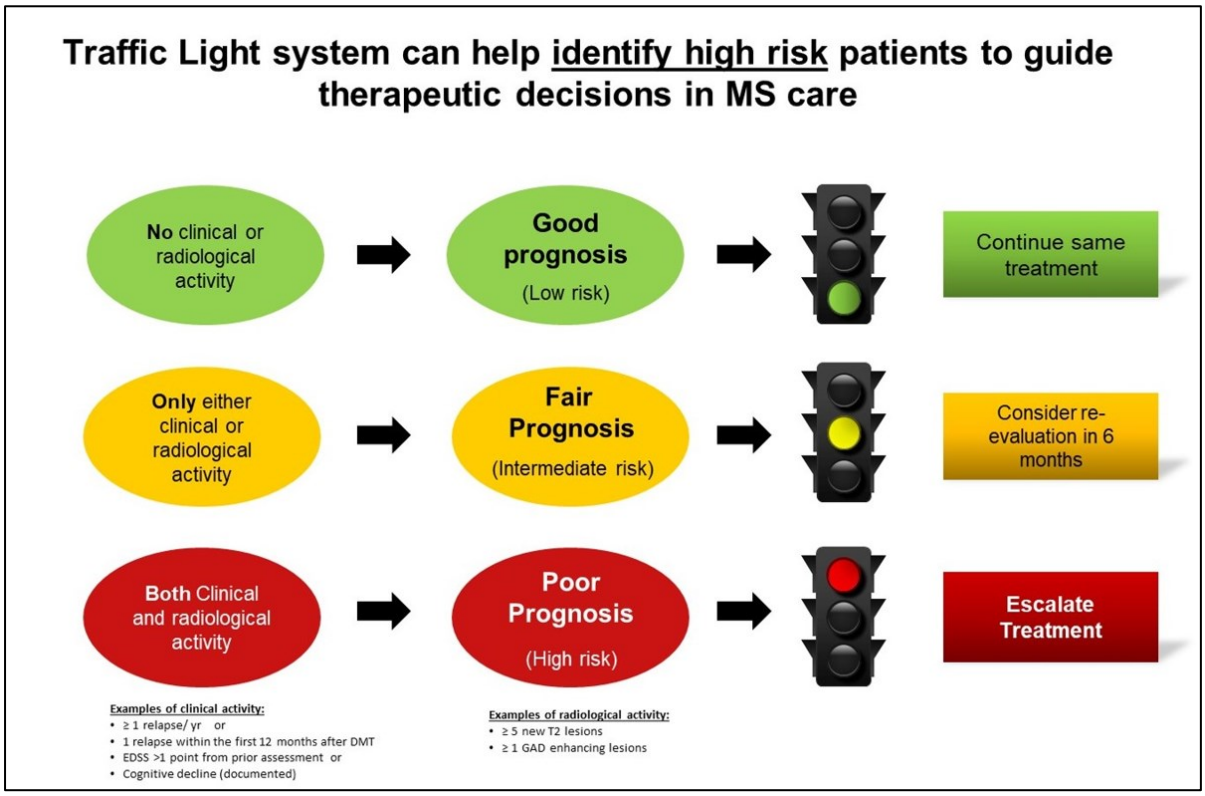
My pilot study (Saposnik, Maurino, et al., 2017) was a double-blind, parallel group, randomised clinical trial. Inclusion criteria included neurologists who are actively involved in managing MS patients. Participants were exposed to 20 simulated case-scenarios (10 cases at baseline, and 10 cases post-randomization to usual care vs. the TLS educational intervention) with and without evidence of MS progression (valence 80%). The educational intervention employed the TLS (See Figure 4 below) to facilitate decisions, allowing participants to easily recognize high-risk scenarios requiring treatment escalation. The TLS consisted of a short, structured, single session intervention of 5 min duration.

The rationale for the design of the TLS was based on the DPT for binary choices in therapeutic decisions. The TLS educational intervention would facilitate optimal therapeutic responses by decreasing automatic responses attributed to system 1 of the DPT and increasing system 2 pathways. For example, simulated case-scenarios showing evidence of disease progression would be associated with the red color of

the TLS, which would trigger a warning sign (meaning “a change is needed“), so participants would retrieve the value of alternative therapeutic options and escalate treatment (See Figure 4 below). In other words, this process would facilitate optimal decisions by requiring participants to make a deliberate choice by retrieving the stored value of different medications for MS and therefore decreasing TI.

**Figure 4: The traffic light system (TLS) educational intervention**

**Panel A.** TLS based algorithm to guide physician’s treatment decisions in MS care




*This figure illustrates how the TLS educational intervention facilitates the decision-making process using traffic light terminology. Participants were exposed to Panel A with a brief introduction of the TLS and then to practice a case-scenario (see Panel B below) to see how this educational intervention would work before starting the simulated-case scenarios. The TLS system creates a link between a color, a risk level, and an action: red light (“high risk”/ “stop and think”), yellow light (“intermediate risk” / “reassess soon”) and green light (“low risk”/”continue the same strategy”). The simulated case-scenario provided updated information about a patient’s current clinical status and evidence of disease progression. This information was coupled with the association of the color “red” and a warning sign (“a change is needed”), thereby facilitating treatment escalation and avoiding the status-quo. As a result, the TLS educational intervention facilitates optimal therapeutic decisions by decreasing automatic responses attributed to system 1 of the DPT and increasing system 2 pathways.*

**Panel B.** Example of simulated case-scenario for the application of the TLS

A 29-year old woman with a diagnosis of RRMS has been on SC interferon beta1a (IFN) for 16 months. She had two recurrent events since the initiation of IFN. Her EDSS score is 2.0. A control brain MRI revealed 6 new T2 bilateral periventricular lesions and one subcortical Gd-enhanced lesion compared to the MRI prior to the initiation of IFN.

**i) Please select the traffic light option that best matches the risk of progression in this case-scenario.**
**Please select <u>one</u>**

**ii) What would you do? Please select <u>one</u>:**
. Stop IFN and start her on Teriflunomide
. Stop IFN and start her on Fingolimod
. Stop IFN and start her on Natalizumab
. Stop IFN and start her on Ocrelizumab
. Stop IFN and start Glatiramer
. Continue IFN and reassess in 6 months

I also measured differences between blocks to invoke decision fatigue. The control group responded as they would do in their usual clinical practice not exposed to the educational intervention. The primary feasibility outcome was the proportion of participants who completed the study and the proportion of participants who correctly identified a high-risk case-scenario with the 'red traffic light'. My target was a 70% accuracy rate in the recognition of the 'red-traffic light'. Secondary outcomes included decision fatigue (defined as an increment of TI in the second block of case-scenarios compared to the first block) and the efficacy of the educational intervention measured as a reduction in TI for MS treatment.

In the larger randomised, controlled trial, 90 participants (neurologists who provide care to MS patients) were randomly assigned to the TLS intervention (n=45) or to the control group (n=45) (Saposnik, Mamdani, et al., 2019). I used the same educational intervention employing the TLS.

The primary outcome was the efficacy of the TLS intervention measured as a reduction in TI. I used the TI score as defined in my pilot study: the number of case scenarios in which a participant showed TI divided by the total number of case scenarios where TI was possible (score ranging from 0 to 1, with higher values representing higher degree of TI). Differences greater than or equal to 0.5 in the TI score were deemed as clinically meaningful (Saposnik, Mamdani, et al., 2019).

Furthermore, participants also had to rate each available medication based on the efficacy and safety profile using a visual analog scale (0 being the lowest and 10 the highest clinical value) representing the contemporary landscape of MS treatment. I created a global score representing the sum of 11 available medications (range: 0-110; with the highest score indicating the highest clinical value). I was interested in assessing the association between the global rating of MS treatments and TI. An overall high score would reflect physicians' favorable perception of the differential benefits of second- and third-line therapies commonly used for treatment escalation. In addition, participants rated a figure representing the current paradigm of treatment escalation (i.e. first, second, and third line therapies) (Figure 1). My goal was to evaluate the association between physician's level of agreement with the current paradigm of treatment escalation (0 being the lowest and 10 the highest agreement with treatment escalation) and the TI score. The proportion of participants showing TI, the associations between the global medication score and figure agreement rate with the TI score were secondary outcome measures.

**Results and discussion:**

In my pilot study, of the 25 participants, 14 were randomly assigned to the control group and 11 to the TLS intervention group. TI was present in 72.0% of participants in at least one case scenario. For the primary feasibility outcome, the completion rate of the study was 100% (25/25 participants). Overall, 77.4% of participants correctly identified the 'red traffic light' for clinical scenarios with high-risk of disease progression. Similarly, 86.4% of participants correctly identified the 'yellow traffic light' for cases that would require a reassessment within 6 to 12 months.

For the secondary fatigue outcome, within-group analysis showed a significant increased prevalence of TI in the second block of case-scenarios (decision fatigue) among participants randomized to the control group (TI pre-intervention 57.1% vs. TI post-intervention 71.4%; p=0.015), but not in the active group (TI pre-intervention 54.6% vs. TI post-intervention 63.6%; p=0.14). For the efficacy outcome, I found a non-significant reduction in TI for the targeted intervention compared to controls (22.6% vs. 33.9% post-intervention; OR 0.57; 95%CI 0.26-1.22). This non-significant difference was expected as this pilot study was underpowered for the efficacy outcome.

I concluded that the TLS educational intervention was feasible and shows some promising results in the identification of high-risk scenarios to reduce decision fatigue and TI. This pilot study provided the basis for the next step to evaluate the efficacy of the TLS in reducing TI.

The larger randomized trial assessing the efficacy of the TLS showed a significant reduction in TI scores in the intervention group (1.47, 95%CI 1.32-1.61) compared to controls (1.93; 95%CI 1.79-2.08); p=0.001. Similarly, the TLS group had a lower prevalence of TI compared to controls (0.67, 95%CI 0.62-0.71 vs. 0.82, 95%CI 0.78-

0.86; p=0.001). The mean global medication score was 68.1 (± 11.0), whereas the mean agreement rate with figure 1 representing the MS landscape was 8.1 (±1.9).

The multivariate analysis, adjusted for age, MS expertise, years of practice, and risk preference showed a significant reduction in TI for participants randomized to the TLS vs control group (β - 0.68, 95%CI: -1.24; -0.11). Specialist status (p=0.002), higher years of experience (p=0.007), and ambiguity aversion (p=0.043) were also associated with lower TI score.

Similarly, the multivariable analysis revealed a 70% reduction in TI after the TLS intervention compared to controls (OR 0.30; 95%CI 0.10-0.89). The adjusted models showed good discrimination (c-statistic=0.74) and calibration (goodness-of-fit test p= 0.52). The TLS educational intervention consistently lowered TI in the intervention group across all TI categories (from participants with low to high TI scores, see Figure 5B of Appendix 5) (Saposnik, Mamdani, et al., 2019).

The analysis of individual responses revealed that for every 100 MS patients with expected disease progression (e.g. both clinical and radiological evidence of disease activity), there will be over 24 patients who will remain with the same treatment if managed by neurologists not exposed to the TLS educational intervention (control group). That number would be decreased to 10 patients if treated by neurologists who received the TLS educational intervention.

The multivariable analysis adjusted for the pre-specified covariates revealed that a higher global medication score was associated with lower TI score (p<0.0001). Similarly, a higher level of agreement with figure 1 representing the landscape of MS for treatment escalation was also associated with lower TI score (p-value<0.0001). These findings suggest that neurologists with a low global score for the MS treatments or those who disagree with the current MS treatment landscape had a significant

higher likelihood of TI. For every 10-point increase in the global medication score, there was a 2.0% reduction in the TI score. For every one-point increase in rating the MS landscape figure, there was a 11.2% reduction in the TI score.

The TLS educational intervention was effective among participants with high and low global medication scores ($p<0.0001$). There was a 70.9% (95%CI 43.7%-98.1%) reduction in the TI score for participants in the high global medication score group and a 41.9% (95%CI 17.2%-66.6%) reduction in the TI score for participants in the low global medication score group. Similarly, my educational intervention was effective in reducing TI ($p<0.0001$) among participants who rated figure 1 representing the landscape above and below the median score. There was a 65.9% (95%CI 44.5%-87.2%) reduction in the TI score for participants who highly rated figure 1 and an 87.8% (95%CI 49.0%-100%) reduction in the TI score for those participants who rated the figure 1 below the median score.

There are few proven effective educational interventions to optimize medical decisions in the literature. A metanalysis of 44 studies comprising 4380 medical professionals showed that cognitive reflection improved diagnostic performance (Prakash et al., 2019). However, most of these studies were not based on a theoretical neuroeconomic framework (G. Elwyn et al., 2011). A Cochrane review showed that decision-aids were among the most effective strategies to bring about more informed, value-based choices, and improved patient-practitioner communication (Stacey et al., 2017). However, there were limited decision-aid tools to assist neurologists in the treatment management of patients with MS.

The TLS offered a unique opportunity to facilitate therapeutic decisions in the medical field, in particular in areas with lacking decision-aid tools such as MS care. For

example, a TLS-based intervention showed the benefits of triaging children with fever by simplifying the decision of hospital admissions based on color-coded risk categories.(M. S. Murphy & Baker, 2014) A similar strategy was used in community mental health by facilitating early assessment of patients with psychosis.(Ashir & Marlowe, 2009)

I concluded that the TLS educational intervention lowered the incidence of TI in MS care irrespective of age, expertise, years for training, and risk preference of participants, which would lead to better patients' outcomes. Participants with low perceptions of benefits of the current MS therapies and those with a lower level of agreement with the contemporary paradigm of treatment escalation had higher TI. The TLS educational intervention was associated with a 68% reduction in the TI score or 70% reduction in the odds of TI, irrespective of participant's specialty, years of practice, beliefs of the benefits of the current MS therapies and the agreement with the paradigm of treatment escalation. In other words, participants exposed to the TLS educational intervention appropriately choose a higher efficacy treatment (e.g. Monoclonal antibodies) instead of the status-quo related to continuing with the same agents (e.g. Glatiramer, Interferon) when evidence of disease progression. The effect of my educational intervention was similar for all categories of TI scores.

**Usability of the TLS:** I also tested the usability of the TLS educational intervention. I surveyed 50 neurologists from Chile, Argentina, and Canada to assess the usability of our TLS educational intervention using the System Usability Score (SUS) (Lewis & Sauro, 2009; Saposnik, Tobler, et al., 2018). The SUS is a validated 10-item questionnaire with five response options. The primary outcome was the average and

95% confidence interval (CI) of the SUS score. Values above 68 are considered highly usable (Bangor, Kortum, & Miller, 2009; Lewis & Sauro, 2009).

*Results and discussion:* the average usability score was 74.7 (95%CI 70.1–79.2). There was one outlier with a score of 35. The usability score excluding the outlier was 76.8 (95%CI 72.7–80.8). The multivariate analysis revealed no association between participants' characteristics and the SUS score. I concluded that my proven effective educational intervention had high usability among neurologists (Saposnik, Tobler, et al., 2018).

## 2.4 Study 4: Autonomic Arousal and Emotional Expressions Mediates Treatment Decisions among Physicians

As mentioned, pupil dilation is a marker of autonomic arousal, which is a reliable proxy for uncertainty (Geng et al., 2015; Lavin et al., 2013; Urai et al., 2017). Thus, pupil dilation might be used as an indicator of how physicians handle uncertainty when making therapeutic decisions. In addition, pupil dilation would also indicate how the TLS educational intervention induces changes in physician’s decisions under uncertainty.

In the fourth group of studies, I evaluated the relationship between arousal (measured by pupil dilation), a proven effective educational intervention, and therapeutic decisions amongst neurologists with expertise in multiple sclerosis (MS). Then, I explored the relationship between emotions and affective states (as captured by muscle facial activity and emotional expressions) and TI in this group of neurologists when making therapeutic decisions.

### 2.4.1 Study on arousal state and TI [Appendix 6]

The main objectives of this study were: i) to evaluate the relation between arousal responses and TI, ii) to investigate how my previously tested and effective educational intervention (Saposnik, Mamdani, et al., 2019) affects arousal responses and TI, and iii) to assess whether arousal responses mediated the association between the educational intervention and TI (i.e. mediation analysis).

**Methods:** In this randomised controlled trial I enrolled 34 neurologists who cared for patients with MS. Participants were randomly assigned to the TLS educational intervention (intervention group, see Figure 4) or usual care (control group). Participants listened to 20 audio-recorded simulated case-scenarios; 16 simulated case-scenarios required treatment escalation and therefore assessed TI, whereas for the remaining four cases, treatment escalation was not recommended. All participants were exposed to the same case-scenarios presented in random order. All simulated case-scenarios were presented auditorily (via headphones connected to the computer) to avoid interference of visual stimulation and automatic eye movements with pupil responses. Participants sat in a room with standardized illumination (see below) with their head fixated by a chin rest and looked at screen while listening to the case-scenarios. Pupil dilation was assessed in time-periods (T) where critical medical information was provided (T1: clinical data, T2: neurological status, T3: MRI data).
The selection of the specific information delivered during those specific time periods was based on the current available evidence from the literature (Rotstein & Montalban, 2019), also supported by our recent worldwide results from a conjoint analysis (manuscript in preparation, data available upon request).

Arousal was measured as the degree of pupil dilation from baseline. Pupil time-series were z-scored to allow comparison of pupil dilation between and within simulated case-scenarios, critical time-periods, and participants. The average pupil size (measured at T0, i.e., 1500ms - 500ms before scenario onset) was taken as pupil baseline (P. R. Murphy et al., 2014). For each simulated case-scenario, I determined arousal responses as the time period-specific maximal z-scored pupil dilation minus the mean baseline z-scored pupil size during T0 (Mathot, 2018; Privitera, Renninger, Carney, Klein, & Aguilar, 2010). Values greater than 0.1 z-scored difference were classified as high arousal periods, following previous research (P. R. Murphy et al., 2014).

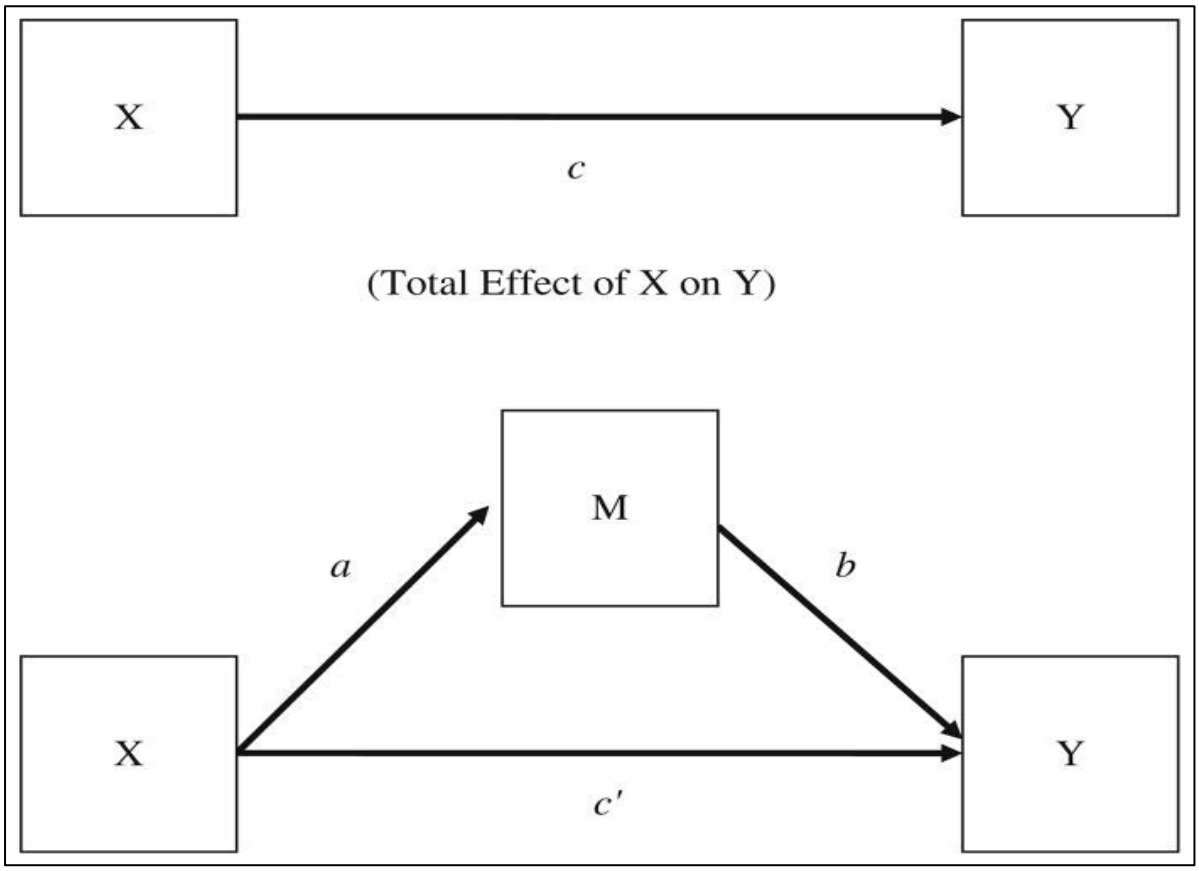
The primary outcome was therapeutic inertia (TI), defined as lack of treatment escalation when warranted based on evidence of disease progression.

Statistical considerations: I applied three analytical approaches: i) comparison of arousal responses across critical time-periods, ii) treatment-effect analysis evaluating the association between arousal responses and TI, and iii) a mediation analysis to assess how the association between individual participant characteristics (e.g.: age, expertise in MS, years of practice, risk preferences) and TI may be mediated by arousal responses. For (i) I used non-parametric tests (Wilcoxon and Mann-Whitney tests for continuous and categorical variables, respectively). For (ii) I compared high vs. low arousal between groups stratified by intervention period (pre- vs. post-intervention). I used generalized estimating equations (GEE) to assess relationships between the variables of interest with TI accounting for clustering (repeated observations on participants). I used the following matrix structure: family (binomial), link (logit), correlation (exchangeable). This analysis was controlled for the pre-defined

explanatory variables, including participant's age, specialist status (MS expert vs. general neurologists), and years of practice, risk preferences, and ambiguity aversion as identified in my previous research (Saposnik, Sempere, et al., 2017). To test the treatment effect of the educational intervention (TLS), I used difference-in-differences models (also called untreated control group design with pre- and post-test) (Antonakis, Bendahan, Jacquart, & Lalive, 2010). This allowed us to measure the treatment effect of the TLS intervention by comparing the change over time (post-test minus pre-test performance) between the intervention and control group. Pupil dilation for each participants and case-scenario was tested as a mediator (see below). As a result, I was able to evaluate whether the benefits of the educational intervention on TI were mediated by individual arousal responsivity.

Mediation analysis (iii) is a technique commonly used in the social sciences and consumer research to explain a relation between an independent variable (e.g. demographic variables) and an outcome via a third variable (called 'mediator') (see below Figure 5) (MacKinnon, Fairchild, & Fritz, 2007; VanderWeele, 2016). The greatest value of mediation analysis in RCT data is that it can establish whether the effects of the intervention (or any independent variable preceding the outcome of interest) on the outcome are mediated by another standardized measured covariate. Here we measured whether the effect of the TLS intervention on TI is mediated by pupil-indexed arousal. The Sobel test determined the significance of the mediation effect, where p-values $<0.05$ indicate a significant mediator (Antonakis et al., 2010).
I used the two-stage least squares (2SLS) regression approach to test for endogeneity due to model misspecification (Antonakis et al., 2010).

**Figure 5: Schematic representation of the relationship between autonomic arousal the TLS educational intervention and therapeutic inertia in MS care**




***Schematic representation of generic mediation analysis.*** *The structural equation model used for the mediation analysis characterizes the relationship between the independent variable X (i.e.: TLS educational intervention) and dependent variables (e.g. age, years of practice, MS expert, risk preference) that are related with an outcome of interest Y (TI). The mediating variable M (pupil enlargement from baseline) is hypothesized to be intermediate in the relation between the educational intervention and TI. This figure illustrates the direct, indirect, and total effect of the educational intervention on TI with pupil dilation as the mediator. Pupil responses mediated the effect of the intervention on TI. The Sobel test confirmed that the mediation effect was significant (p=0.029). The effect of covariates on pupil dilation are omitted in the graph for simplicity.*

**Results and discussion:** Of 38 potentially eligible participants, 34 (89.4%) neurologists completed the study. Before the intervention, there were no differences between groups. TI was present in 50.0% (17/34) of all participants. Pupil dilation was associated with greater TI ($p<0.05$).

i) **Arousal responses predict therapeutic inertia:** I first analyzed differences of pupil dilation by time periods. Overall, pupil size increased for each time period relative to baseline (F-test for linear regression analysis, all p<0.0001). The results remained robust after adjustment for the pre-specified covariates (p<0.0001). Pupil size increased significantly (p<0.0001 adjusted for multiple comparisons, Tukey method) across all ten paired combinations of time-period comparisons (e.g. T1 vs. T0, T1 vs. T2, T1 vs. T3,….T4 vs. T3), except for the comparison between T1 vs T3 (p=0.96).

For every additional standard deviation of pupil dilation, the odds of TI increased by 51% for T1 (95%CI 1.12-2.03), by 31% for T2 (95%CI 1.08-1.59) and by 49% for T3 (95%CI 1.13-1.97). The intervention significantly reduced TI (RR 31.5%; 95%CI 16.1-47.0).

ii) **Time relative to intervention (pre vs. post) and group allocation (intervention vs. control) affect arousal responses:**

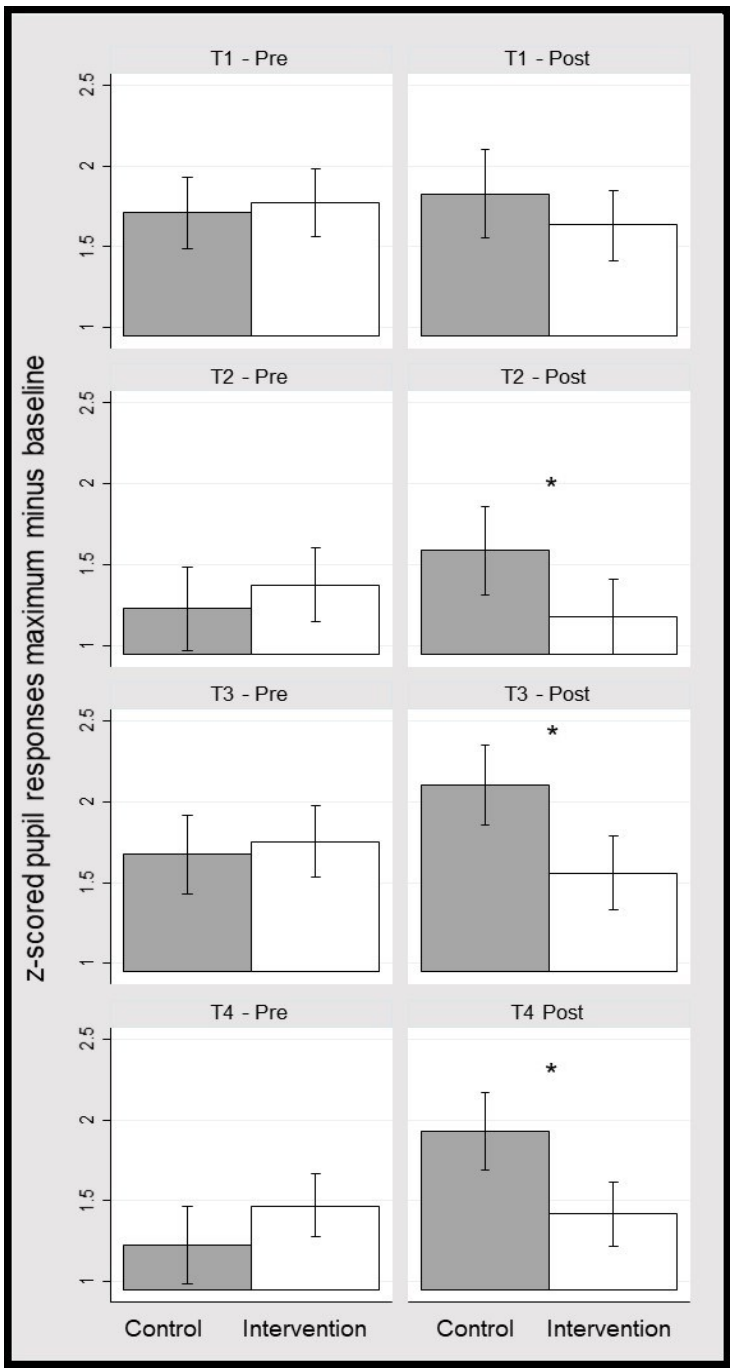
Pupil size did not differ significantly between intervention and control groups before the intervention but did so after the intervention for T2, T3, and T4 (See appendix 6) Overall, the multivariable analysis showed that in the post-intervention period, participants in the control group had significantly enlarged pupils compared to the intervention group for T2 (p=0.049), T3 (p=0.004), and T4 (p<0.0001). No difference was observed for T1 (p=0.47) (see Figure 6 below).

The analysis of dichotomized pupil response (maximum peak minus mean baseline greater than or equal to 0.1 z-scored difference as a high arousal vs. below 0.1 z-scored difference -low arousal response) showed similar results (see Appendix 6).

*Effect of the educational intervention on TI:* I found that participants in the educational intervention group had a significant 31.5% (95%CI 16.1-47.0) reduction in TI

compared to the control group (See Figure 4 in Appendix 7). Linear regression analysis adjusted for participant age, gender, expertise, risk preference and pupil dilation showed that for every therapeutic decision, there was a significant TI decrease of 5.0% (-5.0%, 95%CI -0.8%, -9.3%) in the intervention group. The difference-in-difference analysis revealed no evidence for confounding endogenous effects.

**Figure 6. Effects of intervention and group allocation on pupil responses**




*This figure shows pupil-linked arousal responses (peak minus mean baseline) separately for TLS intervention and control groups, stratified by the intervention period.*

** p<0.01 for the comparison of pupil responses between control and intervention groups.*

*Note that lower responses in the intervention group extend to T4, where no critical information was provided, which may suggest that the protective effect of the intervention extends into the period when participants made decisions under uncertainty*.

**iii) Mediation analysis:** Arousal responses explained 29.0% of the mediated effect of the educational intervention on TI. Other factors (e.g. age, sex, risk preference), had a non-significant or a negligible effect. The sensitivity analyses of adding or removing covariates (i.e., risk preference, age, sex, academic practice) revealed no significant changes in the β coefficients (<10%) of the direct or indirect effects.

The role of autonomic arousal for therapeutic decision making has so far been entirely unexplored. In the present study, I addressed this gap in the framework of therapeutic decisions in MS care, with a focus on decisions not to escalate treatment when recommended by best practice guidelines. I analyzed pupil dilation as an established marker of autonomic arousal (Mathot, 2018; Urai et al., 2017) and found that both continuously measured pupil dilation and dichotomized high vs. low phasic pupil responses were associated with TI. Our data suggest that the intervention may ameliorate TI by reducing arousal responses to critical information. Indeed, pupil dilation mediated the effects of the educational intervention on TI (explaining 29% of the total mediated effect).

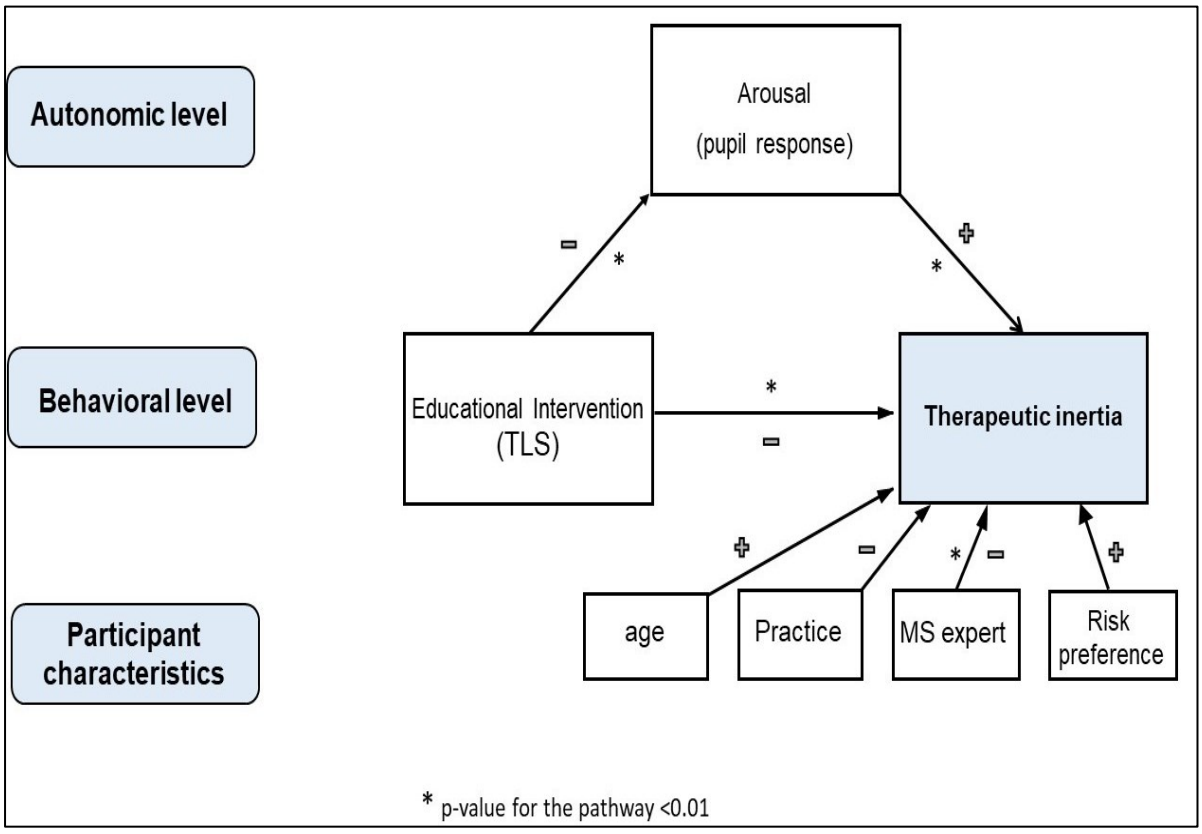
Previous studies have shown that uncertainty drives rapid pupil-linked arousal responses, affecting behavioral choices (Urai et al., 2017) and updating the valuation system based on the presented evidence (de Gee et al., 2014). Phasic pupil size increases are associated with suboptimal or erroneous decisions involving high level of uncertainty (Maier et al., 2019; Wessel, Dolan, & Hollingworth, 2018). Similarly, I presented critical clinical information delivered during the specific time-periods (T1, T2, and T3) simulated MS scenarios to allow participants to update the valuation of each treatment option. My results are consistent with this notion. The TLS educational intervention may reduce arousal by reducing uncertainty and thereby facilitating alternative behavioral strategies. Specifically, the warning function of a red traffic light

highlights the need for switching to a more effective treatment (Saposnik, Mamdani, et al., 2019) and may concurrently boost confidence in the therapeutic decision.

These findings improve our current understanding of the link between pupil dilation as a marker of an autonomic arousal response and physician's decisions. Critical information (i.e., clinical course, functional status and results of recent brain imaging) increases physician's arousal and stronger arousal responses are associated with suboptimal therapeutic decisions (i.e., therapeutic inertia).

Individual arousal responses mediated the treatment effect of an educational intervention and therapeutic decisions in MS care [see Figure 7 below].

**Figure 7.** ***Proposed pathways to explain the relationship between my successful TLS educational intervention and TI***




*This figure illustrates the proposed integration of autonomic, behavioral (decision-making), and participant characteristics associated with TI. The reduction in TI (therapeutic decision) by the educational intervention is mediated by lowering the arousal response. Other covariates may also directly or indirectly influence TI, but in the present study had negligible or non-existent effects (mediation effect <3%).*

Given that I used MS care as a model of complex therapeutic decisions arising in the management of chronic diseases, these findings may also apply to other medical conditions (e.g. hypertension, diabetes, high cholesterol).

### 2.4.2 Study on emotional expressions and TI [Appendix 7]

Emotions play a critical role in our daily decisions (Ekman & Friesen, 2003; Phelps et al., 2014). However, it remains unclear how and what sort of emotional expressions are associated with therapeutic decisions in multiple sclerosis (MS) care.

In this study, I evaluated facial muscle activation (and emotional expression) associated with therapeutic choices, particularly TI. I also sought to evaluate the mediation effect between a physical (e.g. facial muscle activity) or emotional (fear, disgust, surprise) response with a therapeutic decision. Given the known associations between specific facial muscle activation and emotional expression (anger, fear, disgust, surprise, etc.) with an increased attention response that precedes participants' choices (FeldmanHall, Glimcher, Baker, & Phelps, 2016; McDuff, Kodra, Kaliouby, & LaFrance, 2017; Stöckli, Schulte-Mecklenbeck, Borer, & Samson, 2018), I hypothesized that facial muscle activity (e.g. upper lip raise) and emotional expression (disgust, surprise) would increase participants' awareness and therefore mediate the relationship between aversion to ambiguity and TI. I assessed emotional expressions amongst physicians who care for people living with multiple sclerosis (MS) as this care model is representative of the paradigm of complex therapeutic decisions (e.g. multiple therapeutic options with a broad therapeutic range- e.g. different safety and efficacy profiles) in the management of a chronic medical condition.

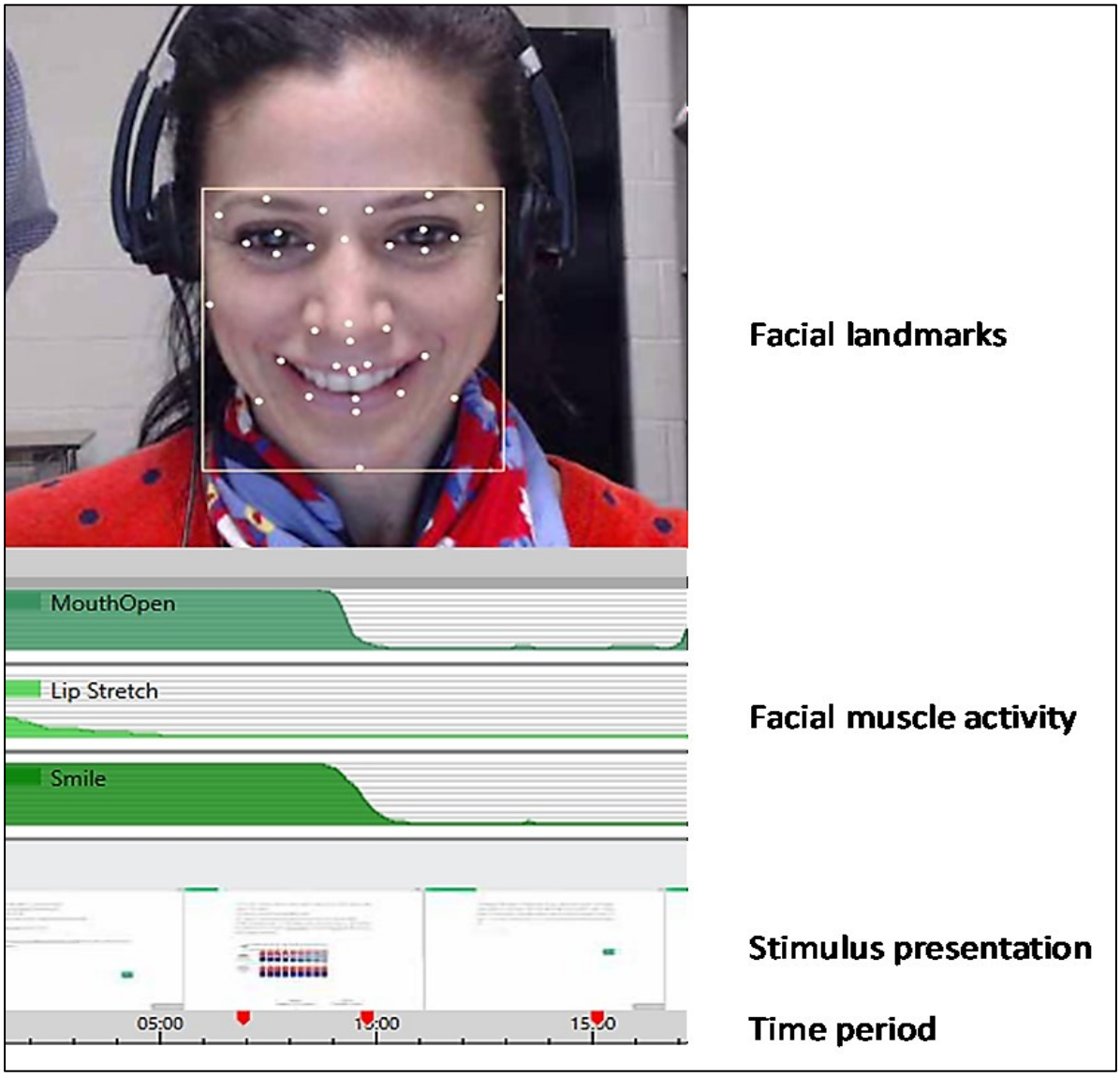
**Methods:** I invited neurologists with expertise in MS to participate in a face-to-face study across Canada. Participants answered questions regarding their clinical practice, aversion to ambiguity, and the management of 10 simulated case-scenarios. I recorded facial muscle activations and their associated emotional expressions during the study, while participants made therapeutic choices.

I used AFFDEX, a machine learning algorithm software that detects emotional expressions based on facial muscle activity (iMotions, 2016; Stöckli et al., 2018). AFFDEX has been validated in more than 7.5 million faces from over 87 countries showing an excellent accuracy (area under the curve greater than 0.9) (https://www.affectiva.com/how/how-it-works/, accessed May 6, 2020). This algorithm uses different features to identify 34 facial landmarks (e.g. eye corners, eye centers, nose tip, mouth corner) with a threshold area, discarding background regions (See Appendix 7).

The region of interest (ROI) contains the whole face including eyes, mouth and nose. AFFDEX applies distinct analytical procedures to identify emotional expressions (https://developer.affectiva.com/mapping-expressions-to-emotions/) (See figure 8 below). During this study, facial detection was recorded to analyze each video frame. Eye blinking and closure were filtered-out. Facial muscle activity is the main unit of study in emotional expressions. Facial movements are detected and mapped on probability values of emotional states (e.g. sadness, joy, disgust, anger, surprise, fear, contempt). The probabilities returned by the AFFDEX module range between zero and one. A value of zero indicates no evidence and a value of one the highest evidence that a certain emotion is fully expressed.(iMotions, 2016) I used raw values of each individual's facial expression to directly compare amongst participants. This approach

mitigates potential errors in the algorithms created to represent emotional expressions due to lack of matching with pre-defined facial muscle activity.

**Figure 8. Facial landmarks for the assessment of facial muscle activity and emotional expressions**




*This figure illustrates the facial landmarks and regions of interest in AFFDEX software. This information is integrated by the iMotions computational software with specific time-points when presented the simulated case-scenarios.*

I used a proxy measure of participants' arousal by combining the level of attention (a summary measure of the timeframe each participant was looking at the screen) and

engagement (a weighted sum of facial expressions). I compared facial muscle activity and emotional expressions between participants with and without TI.

Mixed effects models and mediation analyses were used to evaluate the relationship between ambiguity aversion, facial muscle activity/emotional expressions and TI measured as a binary variable and a continuous score (Saposnik, Oh, et al., 2019).

**Results and discussion:** Of 38 invited participants, 34 (89.4%) neurologists completed the study. The mean age [standard deviation (SD)] was 44.6 (11.5) years; 38.3% were female and 58.8% self-identified as MS specialists. Overall, 17 (50%) participants showed TI in at least one case-scenario and the mean (SD) TI score was 0.74 (0.90). Nineteen (55.9%) participants had aversion to ambiguity in the financial domain. The multivariate analysis adjusted for age, sex and MS expertise showed that aversion to ambiguity in the financial domain (OR 1.56, 95%CI 1.32-1.86) was associated with TI. Most common muscle activations included mouth open (23.4%), brow furrow (20.9%), brow raise (17.6%), and eye widening (13.1%). Most common emotional expressions included fear (5.1%), disgust (3.2%), sadness (2.9%), and surprise (2.8%). After adjustment for age, sex, and physicians' expertise, the multivariate analysis revealed that brow furrow (OR 1.04; 95%CI 1.003-1.09) and lip suck (OR 1.06; 95%CI 1.01-1.11) were associated with an increase in TI prevalence, whereas upper lip raise (OR 0.30; 95%CI 0.15-0.59), and chin raise (OR 0.90; 95%CI 0.83-0.98) were associated with lower likelihood of TI. Disgust and surprise were the emotional expressions associated with a lower TI score (disgust: $p<0.001$; surprise: $p=0.008$) and lower prevalence of TI ($OR_{disgust}$: 0.14, 95%CI 0.03-0.65; $OR_{surprise}$: 0.66, 94%CI 0.47-0.92) after adjusting for covariates. The mediation analysis showed that brow furrow was a partial mediator explaining 21.2% (95%CI 14.9%-38.9%) of the

association between aversion to ambiguity and TI score, followed by nose wrinkle 12.8% (95%CI 8.9%-23.4%). Similarly, disgust was the single emotional expression (partial mediator) that attenuated (-13.2%, 95%CI -9.2% to -24.3%) the effect of aversion to ambiguity on TI.

Previous studies suggest that the neural mechanisms mediating the relation between affect and decisions depend on a participant's emotional arousal and engagement with the specific choice to be made. (Phelps et al., 2014) For example, disgust has been associated with the activation of the insular cortex which may lead to increased arousal modulating the neural responses to aversion to ambiguity, which results in influencing subsequent decision-making (Klucken et al., 2012; Mataix-Cols et al., 2008). Disgust was also shown to increase arousal by modulating emotion-specific attention (van Hooff, van Buuringen, El M'rabet, de Gier, & van Zalingen, 2014). This finding is also consistent with an increased arousal score associated with disgust (and its muscle components) in my study.

These findings are in keeping with this proposed framework as they support an association between facial metrics and emotional expressions (disgust) which may increase participant's awareness/arousal (reflected by increased arousal scores) regarding a compelling decision, thereby showing a reduction in ambiguity aversion and lowering the likelihood of TI. I concluded that facial metrics (e.g. brow furrow, nose wrinkle) and emotional expressions (e.g. disgust) are associated with physicians' choices and partially mediate the effect of aversion to ambiguity on TI (Saposnik, Oh, et al., 2019).

# 3. General Discussion

I focused on five key questions: i) the understanding of most common biases (i.e. status-quo) and personality traits associated with physicians' decisions, ii) factors associated with TI, iii) the role of ambiguity aversion on TI, iv) the development and testing the feasibility and efficacy of my educational intervention using the TLS, and v) the arousal state and emotional expressions associated with therapeutic inertia (TI). In other words, this work comprises a sequence of studies that started with the development of the concept of TI, an outstanding paradigm in daily physicians' decisions. I sought to elucidate factors associated with TI, the roles of arousal responses and emotional states on live treatment decisions in MS care. I started with the conceptualization of TI and ended by identifying determinants and potential mechanisms associated with the likelihood of TI.

## 3.1 Cognitive biases and physician's decisions

The recognition of physicians' cognitive biases are critical to optimize medical decisions, prevent medical errors, provide more realistic patient expectations, and contribute to decreasing the rising health care costs altogether (Andel et al., 2012; van den Berge & Mamede, 2013; Zwaan, Thijs, Wagner, van der Wal, & Timmermans, 2012). I addressed my objectives comprising different medical specialties (anesthesiology, pathology, obstetrics, pediatrics, occupational health, among others) (Baldwin et al., 2005; Crowley et al., 2013; Meyer, Payne, Meeks, Rao, & Singh, 2013; Redelmeier & Shafir, 1995; Saposnik et al., 2013; Stiegler & Ruskin, 2012; Yee et al., 2014). The most critical findings include the high prevalence of cognitive factors and biases (i.e., anchoring and framing effects, information biases) and personality traits

(e.g. tolerance to uncertainty) among physicians, although a precise magnitude is unknown. I was particularly interested in the potential effects of status-quo bias and ambiguity aversion on physician's treatment decisions. There were no studies addressing this question. Studies evaluating physicians' overconfidence, the anchoring effect, and information or availability bias suggest an association with diagnostic inaccuracies (Bytzer, 2007; Crowley et al., 2013; Friedman et al., 2001; Mamede, Schmidt, et al., 2010; Mamede, van Gog, et al., 2010; Meyer et al., 2013; Reyna & Lloyd, 2006). Moreover, anchoring, information bias, overconfidence, premature closure, representativeness and confirmation bias may be associated with therapeutic or management errors (Perneger & Agoritsas, 2011; Redelmeier & Shafir, 1995; Sorum et al., 2003; Stiegler & Ruskin, 2012; Yee et al., 2014). My results were in agreement with a previous systematic review (Blumenthal-Barby & Krieger, 2015). Misinterpretation of recommendations and lower comfort with uncertainty were associated with overutilization of diagnostic tests (e.g. prostate-specific antigen blood test for the diagnosis of prostate cancer in men)(Sorum et al., 2003). Physicians with better coping strategies and tolerance to ambiguity could be related to optimal management (e.g. lower instrumental vaginal deliveries and lower obstetrical errors)(Yee et al., 2014).

I also found sparse data for addressing the relation between physicians' cognitive factors/biases and patient's outcomes. Only one study showed higher complications (OR 1.51, 95%CI 1.10-2.20) among patients cared for by physicians with higher tolerance to ambiguity (Yee et al., 2014). Finally, I identified gaps in the literature as only few (<50%) of an established set of cognitive biases (Croskerry, 2003) were assessed. Other listed and relevant biases were not studied (e.g. status-quo bias, aggregation bias, feedback sanction, hindsight bias). The quality assessment for the

included studies was low to moderate (Saposnik, Redelmeier, et al., 2016). More importantly, only 35% of studies provided information on the association between cognitive biases or personality traits and medical errors (Baldwin et al., 2005; Perneger & Agoritsas, 2011; Redelmeier & Shafir, 1995; Reyna & Lloyd, 2006; Sorum et al., 2003; Stiegler & Ruskin, 2012; Yee et al., 2014).

In conclusion, although cognitive factors/biases may affect a wide range of physicians (and influence diagnostic accuracy, management, and therapeutic decisions), their true prevalence remains unknown. Thus, substantial gaps limit our understanding of the impact of cognitive factors/biases on medical decisions. I proposed the design of more comprehensive studies to evaluate the effect of physicians' personality traits and biases on treatment decisions in live simulated medical encounters. This can be accomplished by identifying physician characteristics, combining validated surveys and experiments commonly used in behavioral economics to elicit several critical personality traits (e.g. tolerance to uncertainty, aversion to risk and ambiguity), and cognitive biases (e.g. status-quo, illusion of control). I believe that this information would provide new insights that may translate into better outcomes (e.g. avoidable hospitalizations, optimized treatment decisions, lower complications related to a procedure or medication, request of unnecessary tests) (Andel et al., 2012; Graber, 2013; Stangierski et al., 2012).

## 3.2 Factors associated with Therapeutic Inertia

Therapeutic decisions are among the most critical tasks made by physicians every day (Gurmankin, Baron, Hershey, & Ubel, 2002). Physicians face the challenge to tailor treatment based on: i) disease activity level, ii) risk of progression, iii) individual patient preferences and characteristics, and iv) personal expertise, in order to identify the optimal balance between safety and efficacy profiles of different medications or interventions with either imperfect or uncertain information (Blumenthal-Barby & Krieger, 2015; Bruck et al., 2013).

My work focused on physicians-level factors associated with TI. Previous studies evaluating the management of different medical conditions (e.g. hypertension, diabetes, chronic pulmonary obstructive disease) suggest that 50 to 70% of clinicians do not intensify therapy when indicated by best practice guidelines (Mohan & Phillips, 2011; O'Connor et al., 2005; Okonofua et al., 2006; L. S. Phillips et al., 2001).

I developed the concept of TI in MS care and conducted a series of studies in different countries (e.g. Spain, Argentina, Chile, Canada) to identify the most common physician-level factors associated with inertia.

Total aversion to ambiguity was observed in one out of four participants in my study (Saposnik, Mamdani, et al., 2019; Saposnik, Sempere, et al., 2017). I showed that ambiguity-averse behavior increased with increasing probability for gains (See Appendix 2, Results study, supplemental file). The most striking finding was that aversion to ambiguity in the financial domain was an independent determinant, increasing the likelihood of TI by at least two-fold. Neurologists with expertise in the care of patients with MS who showed aversion to ambiguity were two- to seven-fold more likely to develop TI after accounting for potential confounders (e.g. age, years of

practice, expertise, number of patients seen per week, practice setting). This is a new contribution to the medical literature.

Although most of my studies showed that ambiguity aversion in the financial domain was associated with an increased likelihood of TI, risk-prone behavior outweighed ambiguity aversion in the Canadian cohort. This finding may be due to the smaller sample size of studies conducted in Canada (Saposnik, Montalban, et al., 2018).

Other factors associated with higher TI included less expertise in MS care, as reflected by either the low number of MS patients seen per week, or by the specialty status (e.g. general neurologists who see patients with MS but are not specialists in this field) (Saposnik, Sempere, et al., 2017). Other studies showed better therapeutic decisions among neurologists with expertise in MS care (Galea et al., 2013; Hobart et al., 2019).

Other factors associated with TI included a physician's low tolerance to uncertainty and the country of practice, with Canada having the lowest TI scores in our research samples (Almusalam et al., 2019; Saposnik, Sempere, et al., 2017). Neurologists with low tolerance to uncertainty had a four-fold higher prevalence of TI (Saposnik, Sempere, et al., 2017). Similar results were found in our larger study cohort (p=0.045) (Almusalam et al., 2019). These findings are also supported by the association between low tolerance to uncertainty and a physician's lower performance in medical residency and among obstetricians (Iannello, Mottini, Tirelli, Riva, & Antonietti, 2017; Yee et al., 2014).

My results have practical implications. I showed that either a simple ambiguity aversion experiment or a tolerance-to-uncertainty survey (without a medical focus) can help to identify TI among neurologists and MS experts with different cultural background, education, practice setting and clinical training (Almusalam et al., 2019). This is

another contribution to the literature. Physician's status quo reflected by the lack of treatment escalation may lead to greater disability of MS patients, increasing the health care costs, and production losses due to incapacity to work. In Europe, the mean annual cost per person with MS has been estimated at €23 000 for EDSS score 0.0–3.5, rising as disability increases to €46 000 for EDSS score 4.0–6.5 and €77 000 for EDSS score 7.0–9.5.(Kobelt, Berg, Lindgren, Fredrikson, & Jonsson, 2006)

Taken together, TI may be explained, at least in part, by: i) the aversion of neurologists to escalate treatment when the available options can have more serious side effects; ii) neurologist's knowledge gaps regarding risk profile of new treatments, and iii) status quo tendency reflected by participants' preference to continue with a known medication profile vs. the unknown risks of a new agent. Other studies have found that TI was associated with lack of training and clinical uncertainty (Kerr et al., 2008). Physicians with better coping strategies and more tolerance to ambiguity may be more likely to choose optimal treatments leading to better patients' outcomes (Yee et al., 2014).

These results lead to my next question: given the magnitude of clinical inertia, can I design an effective educational intervention to overcome the effects of TI?

## 3.3 The Traffic Light System educational intervention ameliorates TI

I started by designing a behavioral-based educational intervention applying the traffic light system (TLS). Each color of the TLS provides a heuristic cue by facilitating the recognition of a difficult situation that prompts a decision (Figure 4). TLS systems couple the color red with a risky situation (e.g. disease progression if treatment is not intensified) that should prompt a behavioral change: the selection of a new medication to intensify treatment. Previous studies using fMRI revealed the existing brain regions

associated with valuation (i.e. ventro-medial prefrontal cortex) and self-control (L. Enax, Krapp, et al., 2015). As such, the TLS would help optimize treatment decisions by interrupting the automatic status-quo state by triggering a re-evaluation processes elicited by the universal warning sign of the color red (Laura Enax et al., 2016). Research studies show optimal food selection when adding color-labels to products with unhealthy components (e.g. high glucose, high sodium, high cholesterol) (Laura Enax et al., 2016; L. Enax, Krapp, et al., 2015; Orriols et al., 2010; X. Zhang et al., 2020). Similar findings were observed in the decision to admit children with fever by guiding pediatricians regarding the course of action (M. S. Murphy & Baker, 2014).

In the pilot study conducted among neurologists from Spain to assess the feasibility of the TLS-based educational intervention, I showed feasibility with a 100% completion rate and promising results with a non-significant 43% reduction in TI post-intervention. These results lead to my larger randomized controlled trial, which assessed the efficacy of the TLS educational intervention in reducing TI among 90 neurologists from Argentina. Although similar (and more complex) strategies to facilitate treatment decisions in MS care exist, our intervention is the only one that has been empirically tested (M. S. Freedman et al., 2013; Stangel et al., 2014).

The TLS educational intervention was associated with 70% reduction in the odds of TI. In other words, participants identified the red traffic light as a risky clinical situation for disease progression and chose a higher efficacy treatment (e.g. Monoclonal antibodies) instead of continuing with the same non-effective medications (e.g. Glatiramer, Interferon). The effect of the TLS educational intervention was similar for all categories of TI scores. My results were robust at the participant level and when analyzing individual responses for both outcome definitions of TI (categories of TI and TI score)(Saposnik, Mamdani, et al., 2019).

There are few effective interventions in medical education that are associated with improvements in clinically meaningful outcomes (Albarqouni, Hoffmann, & Glasziou, 2018). A recent systematic review evaluated 302 controlled studies that investigated the effect of evidence-based educational interventions. Of 85 articles that met the inclusion criteria, 46 (54%) studies were randomised trials, and 51 (60%) included postgraduate level participants. Although the authors evaluated outcomes in multiple domains (e.g. self-efficacy, knowledge, behavior change), none of the studies assessed the benefit to patients (Albarqouni et al., 2018). To the best of my knowledge, the TLS constitutes the first effective educational intervention based on existing brain pathways that is supported by a theoretical framework (e.g.: dual process theory) (Evans, 2003; Tversky & Kahneman, 1974, 1981).

In the evolving landscape of MS treatment, new and more effective agents with improvements in safety profiles are becoming available (Saposnik & Montalban, 2018). Despite such advances, many MS patients remain undertreated (Montalban et al., 2018; Rae-Grant et al., 2018). Our results suggest that the TLS is a useful medical educational intervention, in-keeping with previous research showing better outcomes in the management of obesity, children with fever presenting to emergency rooms, and the selection of healthy food choices (M. S. Murphy & Baker, 2014; Sonnenberg et al., 2013).

In conclusion, I pilot-tested the feasibility and then assessed the efficacy of the TLS educational intervention. I found a significant 70% reduction of TI with our TLS intervention, irrespective of participants medical education, practice settings and perception of the current treatment landscape in MS care. The TLS educational

intervention facilitates optimal therapeutic decisions by decreasing automatic responses (attributed to system 1 of the DPT), and increasing system 2 pathways through detection of the color red (“a change is needed“), facilitating the retrieval and evaluation of alternative therapeutic options, culminating in treatment escalation (moving away from the automatic responses associated with TI).

These findings have practical clinical and health policy implications, which may not only lead to improved patient outcomes, but also lead to the implementation of educational interventions in physicians managing high-risk and complex patients. The TLS intervention has the potential to be translated to other highly prevalent medical conditions, including the management of hypertension, diabetes, and dyslipidemia commonly affecting individuals at high-risk of cardio- and cerebrovascular diseases.

## 3.4 Autonomic Arousal Mediates Biased Treatment Decisions

Upon demonstrating feasibility, usability, and efficacy of the educational intervention, I was interested in understanding how the TLS would reduce TI, including an evaluation of the role of emotional expressions on TI. Several research groups have shown that task-evoked changes in pupil size correlate with different cognitive processes, including conflict processing, surprise, target detection, working memory, attention, and awareness (Geng et al., 2015; Lavin et al., 2013; Preuschoff, t Hart, & Einhauser, 2011; Privitera et al., 2010). Geng et al. showed that evoked-pupil diameter reflects uncertainty during attentional selection, and that uncertainty initiates the involvement of prefrontal cognitive control mechanisms to help disambiguate sensory information and determine the correct response (Geng et al., 2015). Others have shown that uncertainty may trigger rapid pupil-linked arousal responses,

affecting behavioral choices (Urai et al., 2017) and evaluation of newly presented evidence (de Gee et al., 2014). My educational intervention may reduce arousal by decreasing uncertainty about the best course of action, and thereby facilitate alternative decision strategies instead of the status-quo.

I was also cognisant of the influence of emotions on therapeutic decisions, which is largely an unexplored field. To the best of my knowledge (based on a literature review conducted on May 15, 2020), no studies have evaluated facial muscle activation and emotional expression among physicians making decisions. By using a machine learning algorithm software (AFFDEX) that detects emotional expressions based on facial muscle activity, I analyzed facial muscle activation and the emotional expressions of neurologists while they were making therapeutic decisions. I found that emotional expressions (e.g. disgust and surprise) were associated with lower TI. I also recorded which facial manifestations of emotional expressions were associated with TI. Disgust was the single emotion that attenuated the effect of aversion to ambiguity in the financial domain on TI. Disgust has been associated with the activation of the insular cortex, which may lead to increased arousal and modulate the neural response to aversion to ambiguity (influencing subsequent decision-making) (Klucken et al., 2012; Mataix-Cols et al., 2008). Disgust was also shown to increase arousal by modulating emotion-specific attention (van Hooff et al., 2014). These findings were also consistent with an increased arousal score associated with disgust (and its muscle components) in my study.

Traditionally, the striatum, the amygdala, the medial prefrontal, orbitofrontal and insular cortices are thought to process emotional aspects of the decision-making process (Lerner, Li, Valdesolo, & Kassam, 2015; Phelps et al., 2014). Moreover, the

dorsolateral and anterior prefrontal cortices and the posterior parietal cortex may modulate cognitive aspects of decisions (Cohen, 2005). Previous studies have shown that stress reduces activity in the dorsolateral and orbital parts of the prefrontal cortex while enhancing amygdala activity, leading to decreased goal-directed behavior and increased emotional responses (e.g. fear, disgust, contempt) (Otto, Misra, Prasad, & McRae, 2014; Sokol-Hessner, Camerer, & Phelps, 2013). My findings are in keeping with this proposed framework, as they support an association between facial metrics and emotional expressions (disgust), increasing a participant's awareness/arousal, and reducing ambiguity aversion/lowering the likelihood of TI. I hypothesize that the TLS educational intervention decreases the stress associated uncertainty related to treatment options, thus increasing goal-directed behavior (i.e. treatment escalation) and decreasing emotional responses (e.g. fear, disgust, contempt) associated with TI. Taken together, these findings suggest that interventions reducing TI may partly rely on emotional factors (Saposnik, Maurino, et al., 2017) and arousal responses, which both play a more important role in medical decision-making than hitherto assumed.

Given the limited training most physicians undergo in risk management and medical decision-making, my results support the notion that physicians are vulnerable when handling decisions under uncertainty, especially if they have an aversion to ambiguity (Dijkstra, Pols, Remmelts, & Brand, 2015; Kostopoulou, Russo, Keenan, Delaney, & Douiri, 2012; Monrouxe et al., 2017).

These findings also have practical clinical implications. Using the autonomic arousal response as a marker of effectiveness, I was able to identify 25% of physicians making suboptimal decisions and estimate the inertia-reducing benefits of novel educational interventions. Consequently, these results open avenues to tailor educational interventions and formal risk-assessment training to decision makers (medical

students, family doctors, and specialists) and conditions. In the future, this approach may help optimize therapeutic decisions for other more prevalent chronic diseases (e.g. hypertension, diabetes) leading to better clinical outcomes and improved medical education.

In summary, the critical medical information commonly used by physicians to make complex treatment decisions (i.e., clinical course, functional status of the patient and results of recent brain imaging), triggers an increased arousal response, which are associated with suboptimal therapeutic decisions (i.e., therapeutic inertia). Moreover, the inertia-reducing effects of an educational intervention appear to be mediated by a reduction in the arousal state. The second study confirms our findings from two observational studies (Saposnik, Mamdani, et al., 2019; Saposnik, Sempere, et al., 2017) that ambiguity aversion is associated with an increased TI (Saposnik, Oh, et al., 2019). I showed an association between emotional expressions (i.e. disgust), ambiguity aversion, and TI. For example, disgust may reduce the effects of ambiguity aversion on TI by decreasing the expression of ambiguity aversion.

## 3.5 Limitations

The described studies have several limitations that deserve comment. First, although my aimed was to be as systematic as possible in reviewing the literature, I cannot rule out involuntary omissions (Saposnik, Redelmeier, et al., 2016). I was not able to conduct a meta-analysis due to significant differences in study design, diversity of definitions and data reported, outcome measures, and the overall small number of studies evaluating specific cognitive factors or biases. Second, as described in the literature most of the studies were based on simulated case-scenarios, which may not truly reflect therapeutic decisions in clinical practice. Third, some of my studies

(Saposnik, Maurino, et al., 2017; Saposnik, Oh, et al., 2019) had a small sample size, which may affect the precision of estimated coefficients and the generalizability of the results. Fourth, the variability of participant's characteristics, practice settings, and level of education may explain some disparities when cross-comparing studies conducted in different countries (Almusalam et al., 2019; Saposnik, Sempere, et al., 2017). Fifth, I have no information regarding the long-term benefits of the TLS education intervention on TI given the study design. Future research will be needed to investigate this question. Sixth, as mentioned, most of studies included in this thesis do not consider the concept of shared decision-making. However, shared decisions can only take place once physicians identify the best course of action before discussing appropriate treatment options with patients. Seventh, pupil size is not a standard measure in clinical practice. As a result, my study should be interpreted as a proof-of-concept, suggesting that there is value in testing this metric in future studies. Eighth, arousal responses may have different triggers and effects than the ones I tested and be dependent on sensory modality (e.g. auditory vs. visual presentation of case-scenarios) and educational strategy. Furthermore, I tested arousal responses related to our designed TLS intervention, which may differ from autonomic responses associated with other educational interventions or tested by other methods (i.e.; galvanic skin responses, heart rate variability). Ninth, despite the extensive validation of the AFFDEX software, the association between emotional states and TI may require further assessment given the high variability of muscle group activations that are mapped to emotional expressions. It is possible that other emotions may influence TI but were not detected in our study due to lack of power. As such, my results should be interpreted with caution considering there are no other similar studies available for comparison. Despite these limitations, my studies provide critical information on the

foundations of therapeutic decisions among practicing physicians, how they handle uncertainty and a proven strategy to overcome TI.

## 3.6 Future Directions

My work provides critical information on the foundations of therapeutic inertia. By elucidating cognitive biases and factors associated with TI (and developing an effective educational intervention), my studies open opportunities for investigation in behavioral sciences, clinical research, guideline-based decisions, and medical education.(Schwartz, 2011, 2013) For example, further studies may facilitate the identification of physicians at high risk of having TI by evaluating physical, emotional, and behavioral responses, such that educational interventions can be tailored to these individuals. These results also raise further questions in translational research. For example, future studies using fMRI and computational modelling would provide information regarding the location of encoded values for therapeutic decisions and the underlying brain pathways (e.g. involvement of the DLPFC, vmPFC, locus coeruleus and amygdala) associated with TI. For example, I would evaluate the specific brain pathways associated with the implementation of the TLS system using fMRI. My work on therapeutic inertia provides the foundations for future research in this area.

Another crucial next step would be to translate this knowledge into the development of programs in medical education. For example, the addition of new courses that integrate risk management strategies with the learned foundations on therapeutic decisions in graduate and postgraduate medical education could foster more optimal decision-making and overcome current clinical gaps.

Another consideration is the application of biosensors and wearable devices already implemented in patient care for diagnostic and therapeutic decisions (Carroll,

Kobylecki, Silverdale, Thomas, & group, 2019; Lambert, Bumgarner, & Tarakji, 2019; Lown et al., 2020; Semaan et al., 2020). Advances in eye-tracking and biosensor technologies, machine learning, and artificial intelligence may lead to the creation of hardware and software to warn of suboptimal decisions by relying on live autonomic arousal responses in routine clinical practice. Although this idea may not be feasible at the present time, developing technology that incorporates augmented reality using smart glasses or wearable biosensors (tested in consumer research; (Muhammad Sayem, Hon Teay, Shahariar, Fink, & Albarbar, 2020; Peake, Kerr, & Sullivan, 2018)) could facilitate the application of color-coded warning signs (either based on pupil or heart rate variability as markers of central arousal) to assist physicians with diagnostic and therapeutic decisions (Shi, Zhao, Liu, Gao, & Dou, 2019). Given the unconscious process of autonomic responses (triggered by new information presented in medical encounters), which usually precedes clinical decisions, this approach would bring another layer of safety to prevent medical errors.

I expect that the implementation of these strategies could result in better medical decisions and result in improved patient outcomes (e.g. lower disability, better quality of life), more effective and patient-oriented clinical encounters, and lower health care costs.

# 4. General Conclusion

- Making decisions in medical care is a difficult task involving a variety of cognitive processes (Glimcher & Fehr, 2014). Decisions based on erroneous assessments may result in unrealistic patient and family expectations, suboptimal advice, and incorrect treatment. Suboptimal decision making and medical errors are exemplified by the concept of therapeutic inertia (TI).
- How doctors handle uncertainty (e.g. having imperfect information, an unknown treatment response for a particular patient, personal cognitive biases or personality traits) is highly relevant to therapeutic decision-making (Kerr et al., 2008; Levy et al., 2010; Mohan & Phillips, 2011; Reach, 2014; Saposnik, Mamdani, et al., 2019; Saposnik, Sempere, et al., 2017; Saposnik, Sempere, et al., 2016a).
- Given that most therapeutic choices are binary, I used the framework provided by the DPT to explain medical decisions under uncertainty, incorporating physician's cognitive biases, risk preferences, and ambiguity aversion.
- Multiple sclerosis is a paradigm of complex treatment decisions due to the increasingly broad landscape of medications with different efficacy and safety profiles, which makes it an ideal condition in which to evaluate TI.
- The central question of this thesis revolves around how physicians handle uncertainty and how they can avoid making suboptimal decisions. I broke down this problem into five specific inquiries with the following conclusions:
    - Are cognitive biases common? Our systematic review found that cognitive biases and personality traits affect over 50% of practicing physicians across different specialties and highlighted literature gaps related to the impact of status-quo and aversion to ambiguity on treatment decisions.

- What is the prevalence and what are the most common factors associated with TI? How does physician's aversion to ambiguity and low tolerance to uncertainty influence their treatment choices? I found that TI is a common phenomenon affecting over 60% of physicians who care for patients with different chronic medical conditions, and 60-90% of MS neurologists. TI is evident in nearly one out of four treatment decisions with correspondingly negative consequences for patients and the health care system (Almusalam et al., 2019; Saposnik, Mamdani, et al., 2019; Saposnik, Sempere, et al., 2017). Neurologists' aversion to ambiguity, low tolerance to uncertainty, and expertise are the most common physician-level determinants of TI.
- What can be done to ameliorate the effects of TI? Our innovative TLS educational intervention was designed to take advantage of an existing learned pathway that couples the color red with a learned warning sign, and prevents automatic responses related to the status-quo (no treatment change when warranted by best practice recommendations) (L. Enax, Krapp, et al., 2015; Saposnik, Mamdani, et al., 2019; X. Zhang et al., 2020). In a randomized controlled trial, our TLS educational intervention decreased the likelihood of TI by 70% (Saposnik, Mamdani, et al., 2019). The effects of the TLS were independent of demographic factors, physician's expertise, practice setting, and different perceptions about the current available treatment.
- Physician's favorable ratings of the benefits of MS medications (measured by the global medication score) and level of agreement with the paradigm with treatment escalation were associated with lower TI.
- The TLS educational intervention would facilitate optimal therapeutic responses by decreasing automatic responses attributed to system 1 of the DPT, and

increasing the following system 2 pathways: a simulated case-scenario showing evidence of disease progression → identification of red color of the TLS → triggers a warning sign ("a change is needed") → retrieval of the value associated with alternative therapeutic options → treatment escalation (moving away from continuation with the same treatment as represented by TI).

- What is the relationship between our successful TLS educational intervention, arousal responses, emotional expressions, and TI? I observed that i) critical medical information provided in specific time-periods during the presentation of simulated case-scenarios elicited a phasic arousal response across all participants in the pre-intervention period, ii) the control group had a significant pupil enlargement in the post-intervention period when comparing TLS intervention groups for most critical time periods, iii) despite the limited power of this study, the difference-in-differences analysis revealed a significant 31.5% (95%CI 16.1-47.0) reduction in TI if neurologists were exposed to the educational intervention, and iv) arousal responses regulate the treatment effect of our TLS educational intervention and therapeutic decisions in MS care.
- The role of emotions in decision-making cannot be ignored. When applying a validated machine learning algorithm, disgust and surprise were associated with a lower TI. It is possible that other emotions that were less frequently expressed in our study also play a role in decreasing TI.

- The results of my studies suggest that pupil dilation, a marker of central arousal, is an indicator of how physicians handle uncertainty when making therapeutic decisions. The arousal response also mediates how the TLS-based educational intervention changes physician's treatment decisions under uncertainty, thus decreasing TI.

- In other words, the TLS educational intervention decreases the stress associated with uncertainty related to treatment options, thus increasing goal-directed behavior (i.e. treatment escalation) and decreasing emotional responses (e.g. fear, disgust, contempt) associated with TI.
- An optimal treatment choice is based on the assessment of the benefits and side effects (risks) of a specific medication, with the overarching goal of patients achieving the best possible outcome for patients. The decision to continue with the same medication (i.e. TI) or escalate to a more efficacious treatment is based on computations of medications value commonly influence by ambiguity aversion, physician's expertise, and regulatory aspects of the country of practice.
- The findings of my studies have practical clinical implications in patient care, post-graduate medical education, and health policy. We are now able to identify how physicians manage uncertainty when making treatment decisions. We can estimate the inertia-reducing benefits of our novel TLS educational intervention and use the autonomic arousal response as a marker of their effectiveness that is unaffected by demand effects or cognitive biases. Consequently, these findings open avenues to tailor educational interventions and formal risk-assessment training to decision makers (medical students, family doctors, and specialists) and conditions. Furthermore, the public, patients and their families may demand a formal training in medical decision-making (e.g. increased awareness of biases, foundations of TI, decision-theory) of future doctors enrolled at medical schools. In the future, this approach may help optimize long-term therapeutic decisions for other and more prevalent chronic diseases (e.g. hypertension, diabetes, high cholesterol) leading to better clinical outcomes by improving medical education.

- As a corollary of my studies, physicians should update their evaluation of treatment options based on new clinical information provided by simulated MS scenarios.
- I hope that these results increase awareness among health care providers, patients and their families, and policymakers to recognize the importance of identifying factors influencing physician's decisions, understanding how they handle uncertainty and implementing strategies to ameliorate TI and avoid medical errors.

# References


Aberegg, S. K., Haponik, E. F., & Terry, P. B. (2005). Omission bias and decision making in pulmonary and critical care medicine. *Chest, 128*(3), 1497-1505. doi:10.1378/chest.128.3.1497

Ahmed, M., Oh, A., Vanderlee, L., Franco-Arellano, B., Schermel, A., Lou, W., & L'Abbé, M. R. (2020). A randomized controlled trial examining consumers' perceptions and opinions on using different versions of a FoodFlip© smartphone application for delivery of nutrition information. *Int J Behav Nutr Phys Act, 17*(1), 22. doi:10.1186/s12966-020-0923-1

Albarqouni, L., Hoffmann, T., & Glasziou, P. (2018). Evidence-based practice educational intervention studies: a systematic review of what is taught and how it is measured. *BMC Med Educ, 18*(1), 177. doi:10.1186/s12909-018-1284-1

Allison, J. J., Kiefe, C. I., Cook, E. F., Gerrity, M. S., Orav, E. J., & Centor, R. (1998). The association of physician attitudes about uncertainty and risk taking with resource use in a Medicare HMO. *Med Decis Making, 18*(3), 320-329. doi:10.1177/0272989x9801800310

Almusalam, N., Oh, J., Terzaghi, M., Maurino, J., Bakdache, F., Montoya, A., . . . Saposnik, G. (2019). Comparison of Physician Therapeutic Inertia for Management of Patients With Multiple Sclerosis in Canada, Argentina, Chile, and Spain. *JAMA Netw Open, 2*(7), e197093. doi:10.1001/jamanetworkopen.2019.7093

Andel, C., Davidow, S. L., Hollander, M., & Moreno, D. A. (2012). The economics of health care quality and medical errors. *J Health Care Finance, 39*(1), 39-50.

Anderson, L. R., & Mellor, J. M. (2008). Predicting health behaviors with an experimental measure of risk preference. *J Health Econ, 27*(5), 1260-1274. doi:10.1016/j.jhealeco.2008.05.011

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly, 21*(6), 1086-1120. doi:https://doi.org/10.1016/j.leaqua.2010.10.010

Ashir, M., & Marlowe, K. (2009). Traffic Lights: A practical risk management system for community early intervention in psychosis teams. *Clinical Governance: An International Journal, 14*(3), 226-235. doi:10.1108/14777270910976166

Bagnjuk, J., Konig, H. H., & Hajek, A. (2019). Personality Traits and Obesity. *Int J Environ Res Public Health, 16*(15). doi:10.3390/ijerph16152675

Baldwin, R. L., Green, J. W., Shaw, J. L., Simpson, D. D., Bird, T. M., Cleves, M. A., & Robbins, J. M. (2005). Physician risk attitudes and hospitalization of infants with bronchiolitis. *Acad Emerg Med, 12*(2), 142-146. doi:10.1197/j.aem.2004.10.002

Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: adding an adjective rating scale. *J. Usability Studies, 4*(3), 114-123.

Bates, D. W., Cullen, D. J., Laird, N., Petersen, L. A., Small, S. D., Servi, D., . . . et al. (1995). Incidence of adverse drug events and potential adverse drug events. Implications for prevention. ADE Prevention Study Group. *JAMA, 274*(1), 29-34.

Bekker, H., Thornton, J. G., Airey, C. M., Connelly, J. B., Hewison, J., Robinson, M. B., . . . Pearman, A. D. (1999). Informed decision making: an annotated bibliography and systematic review. *Health Technol Assess, 3*(1), 1-156.

Bermel, R. A., You, X., Foulds, P., Hyde, R., Simon, J. H., Fisher, E., & Rudick, R. A. (2013). Predictors of long-term outcome in multiple sclerosis patients treated with interferon beta. *Ann Neurol, 73*(1), 95-103. doi:10.1002/ana.23758

Binmore, K., Stewart, L., & Voorhoeve, A. (2012). How Much Ambiguity Aversion? Finding Indifferences between Ellsberg's Risky and Ambiguous Bets. *Journal of Risk and Uncertainty, 45*(3), 215-238.

Blumenthal-Barby, J. S., & Krieger, H. (2015). Cognitive biases and heuristics in medical decision making: a critical review using a systematic search strategy. *Med Decis Making, 35*(4), 539-557. doi:10.1177/0272989X14547740

Bruck, W., Gold, R., Lund, B. T., Oreja-Guevara, C., Prat, A., Spencer, C. M., . . . Zamvil, S. S. (2013). Therapeutic decisions in multiple sclerosis: moving beyond efficacy. *JAMA Neurol, 70*(10), 1315-1324. doi:10.1001/jamaneurol.2013.3510

Burks, J., Marshall, T. S., & Ye, X. (2017). Adherence to disease-modifying therapies and its impact on relapse, health resource utilization, and costs among patients with multiple sclerosis. *ClinicoEconomics and outcomes research : CEOR, 9*, 251-260. doi:10.2147/CEOR.S130334

Bytzer, P. (2007). Information bias in endoscopic assessment. *Am J Gastroenterol, 102*(8), 1585-1587. doi:10.1111/j.1572-0241.2006.00911.x

Camerer, C., & Weber, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty, 5*(4), 325-370. doi:10.1007/bf00122575

Caplin, A., & Glimcher, P. W. (2014). Chapter 1 - Basic Methods from Neoclassical Economics. In P. W. Glimcher & E. Fehr (Eds.), *Neuroeconomics (Second Edition)* (pp. 3-17). San Diego: Academic Press.

Carney, P. A., Yi, J. P., Abraham, L. A., Miglioretti, D. L., Aiello, E. J., Gerrity, M. S., . . . Elmore, J. G. (2007). Reactions to uncertainty and the accuracy of diagnostic mammography. *J Gen Intern Med, 22*(2), 234-241. doi:10.1007/s11606-006-0036-9

Carroll, C., Kobylecki, C., Silverdale, M., Thomas, C., & group, P. K. G. a. (2019). Impact of Quantitative Assessment of Parkinson's Disease-Associated Symptoms Using Wearable Technology on Treatment Decisions. *J Parkinsons Dis, 9*(3), 601. doi:10.3233/JPD-191623

Carter, C. S., & van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cogn Affect Behav Neurosci, 7*(4), 367-379. doi:10.3758/cabn.7.4.367

Christopoulos, G. I., Tobler, P. N., Bossaerts, P., Dolan, R. J., & Schultz, W. (2009). Neural correlates of value, risk, and risk aversion contributing to decision making under risk. *J Neurosci, 29*(40), 12574-12583. doi:10.1523/JNEUROSCI.2614-09.2009

Classen, D. C., Pestotnik, S. L., Evans, R. S., & Burke, J. P. (1991). Computerized surveillance of adverse drug events in hospital patients. *JAMA, 266*(20), 2847-2851.

Clewett, D. V., Huang, R., Velasco, R., Lee, T.-H., & Mather, M. (2018). Locus coeruleus activity strengthens prioritized memories under arousal. *Journal of Neuroscience, 38*(6), 1558-1574.

Cohen, J. D. (2005). The Vulcanization of the Human Brain: A Neural Perspective on Interactions Between Cognition and Emotion. *Journal of Economic Perspectives, 19*(4), 3-24. doi:doi: 10.1257/089533005775196750

Cooke, C. E., Sidel, M., Belletti, D. A., & Fuhlbrigge, A. L. (2012). Review: clinical inertia in the management of chronic obstructive pulmonary disease. *COPD, 9*(1), 73-80. doi:10.3109/15412555.2011.631957

Corathers, S. D., & DeSalvo, D. J. (2020). Therapeutic Inertia in Pediatric Diabetes: Challenges to and Strategies for Overcoming Acceptance of the Status Quo. *Diabetes Spectr, 33*(1), 22-30. doi:10.2337/ds19-0017

Cristiano, E., Alonso, R., Alvez Pinheiro, A., Bacile, E. A., Balbuena, M. E., Ballario, C., . . . Correale, J. (2018). Argentinean recommendations on the identification of treatment

failure in relapsing remitting multiple sclerosis patients. *J Neurol Sci, 385*, 217-224. doi:10.1016/j.jns.2018.01.004

Croskerry, P. (2003). The importance of cognitive errors in diagnosis and strategies to minimize them. *Acad Med, 78*(8), 775-780.

Crowley, R. S., Legowski, E., Medvedeva, O., Reitmeyer, K., Tseytlin, E., Castine, M., . . . Mello-Thoms, C. (2013). Automated detection of heuristics and biases among pathologists in a computer-based system. *Adv Health Sci Educ Theory Pract, 18*(3), 343-363. doi:10.1007/s10459-012-9374-z

Cunningham, B. A., Bonham, V. L., Sellers, S. L., Yeh, H. C., & Cooper, L. A. (2014). Physicians' anxiety due to uncertainty and use of race in medical decision making. *Med Care, 52*(8), 728-733. doi:10.1097/MLR.0000000000000157

Davis, E. L., McCaffery, K., Mullan, B., & Juraskova, I. (2015). An exploration of decision aid effectiveness: the impact of promoting affective vs. deliberative processing on a health-related decision. *Health Expect, 18*(6), 2742-2752. doi:10.1111/hex.12248

de Gee, J. W., Knapen, T., & Donner, T. H. (2014). Decision-related pupil dilation reflects upcoming choice and individual bias. *Proc Natl Acad Sci U S A, 111*(5), E618-625. doi:10.1073/pnas.1317557111

Diaz Rodriguez, A., Murga, N., Camafort-Babkowski, M., Lopez Peral, J. C., Ruiz, E., Ruiz-Baena, J., & Valdivielso, P. (2014). Therapeutic inertia in hypercholesterolaemia is associated with ischaemic events in primary care patients. A case-control study. *Int J Clin Pract, 68*(8), 1001-1009. doi:10.1111/ijcp.12419

Dijkstra, I. S., Pols, J., Remmelts, P., & Brand, P. L. (2015). Preparedness for practice: a systematic cross-specialty evaluation of the alignment between postgraduate medical education and independent practice. *Med Teach, 37*(2), 153-161. doi:10.3109/0142159X.2014.929646

Dijkstra, I. S., Pols, J., Remmelts, P., Rietzschel, E. F., Cohen-Schotanus, J., & Brand, P. L. (2015). How educational innovations and attention to competencies in postgraduate medical education relate to preparedness for practice: the key role of the learning environment. *Perspect Med Educ, 4*(6), 300-307. doi:10.1007/s40037-015-0219-3

Djulbegovic, B., & Elqayam, S. (2017). Many faces of rationality: Implications of the great rationality debate for clinical decision-making. *J Eval Clin Pract, 23*(5), 915-922. doi:10.1111/jep.12788

Djulbegovic, B., Elqayam, S., & Dale, W. (2018). Rational decision making in medicine: Implications for overuse and underuse. *J Eval Clin Pract, 24*(3), 655-665. doi:10.1111/jep.12851

Djulbegovic, B., Hozo, I., Beckstead, J., Tsalatsanis, A., & Pauker, S. G. (2012). Dual processing model of medical decision-making. *BMC Med Inform Decis Mak, 12*, 94. doi:10.1186/1472-6947-12-94

Djulbegovic, B., Hozo, I., Mayrhofer, T., van den Ende, J., & Guyatt, G. (2019). The threshold model revisited. *Journal of Evaluation in Clinical Practice, 25*(2), 186-195. doi:10.1111/jep.13091

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences.

Duquette, P., Giacomini, P. S., Bhan, V., Hohol, M., & Schecter, R. (2016). Balancing Early Aggression Against Risk of Progression in Multiple Sclerosis. *Can J Neurol Sci, 43*(1), 33-43. doi:10.1017/cjn.2015.302

Ekman, P., & Friesen, W. V. (2003). *Unmasking the face: A guide to recognizing emotions from facial clues*: Ishk.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics, 75*(4), 643-669. doi:10.2307/1884324

Elstein, A. S., & Schwartz, A. (2002). Clinical problem solving and diagnostic decision making: selective review of the cognitive literature. *BMJ, 324*(7339), 729-732.

Elwyn, G., Frosch, D., Thomson, R., Joseph-Williams, N., Lloyd, A., Kinnersley, P., . . . Barry, M. (2012). Shared Decision Making: A Model for Clinical Practice. *Journal of General Internal Medicine, 27*(10), 1361-1367. doi:10.1007/s11606-012-2077-6

Elwyn, G., Stiel, M., Durand, M. A., & Boivin, J. (2011). The design of patient decision support interventions: addressing the theory-practice gap. *J Eval Clin Pract, 17*(4), 565-574. doi:10.1111/j.1365-2753.2010.01517.x

Ely, J. W., Graber, M. L., & Croskerry, P. (2011). Checklists to reduce diagnostic errors. *Acad Med, 86*(3), 307-313. doi:10.1097/ACM.0b013e31820824cd

Enax, L., Hu, Y., Trautner, P., & Weber, B. (2015). Nutrition labels influence value computation of food products in the ventromedial prefrontal cortex. *Obesity (Silver Spring), 23*(4), 786-792. doi:10.1002/oby.21027

Enax, L., Krajbich, I., & Weber, B. (2016). Salient nutrition labels increase the integration of health attributes in food decision-making. *Judgment and Decision Making, 11*(5), 460-471.

Enax, L., Krapp, V., Piehl, A., & Weber, B. (2015). Effects of social sustainability signaling on neural valuation signals and taste-experience of food products. *Front Behav Neurosci, 9*, 247. doi:10.3389/fnbeh.2015.00247

Ernstsson, O., Gyllensten, H., Alexanderson, K., Tinghög, P., Friberg, E., & Norlund, A. (2016). Cost of Illness of Multiple Sclerosis - A Systematic Review. *PLoS One, 11*(7), e0159129-e0159129. doi:10.1371/journal.pone.0159129

Evans, J. S. (2003). In two minds: dual-process accounts of reasoning. *Trends in cognitive sciences, 7*(10), 454-459. doi:10.1016/j.tics.2003.08.012

Fargen, K. M., & Friedman, W. A. (2014). The science of medical decision making: neurosurgery, errors, and personal cognitive strategies for improving quality of care. *World Neurosurg, 82*(1-2), e21-29. doi:10.1016/j.wneu.2014.03.030

Fargen, K. M., Leslie-Mazwi, T. M., Chen, M., & Hirsch, J. A. (2020). Physician, know thyself: implicit and explicit decision-making for mechanical thrombectomy in stroke. *J Neurointerv Surg*. doi:10.1136/neurintsurg-2020-015973

Feinstein, A., Freeman, J., & Lo, A. C. (2015). Treatment of progressive multiple sclerosis: what works, what does not, and what is needed. *Lancet Neurol, 14*(2), 194-207. doi:10.1016/S1474-4422(14)70231-5

FeldmanHall, O., Glimcher, P., Baker, A. L., & Phelps, E. A. (2016). Emotion and decision-making under uncertainty: Physiological arousal predicts increased gambling during ambiguity but not risk. *Journal of experimental psychology. General, 145*(10), 1255-1262. doi:10.1037/xge0000205

Fleming, S. M., Thomas, C. L., & Dolan, R. J. (2010). Overcoming status quo bias in the human brain. *Proceedings of the National Academy of Sciences, 107*(13), 6005-6009. doi:10.1073/pnas.0910380107

Flexner, A. (2002). Medical education in the United States and Canada. *Bulletin of the World Health Organization, 80*, 594-602.

Freedman, M. S., Duquette, P., Grand'Maison, F., Lee, L., Vorobeychik, G., Lara, N., . . . Haddad, P. (2019). The clinical and cost impact of switching to fingolimod versus other first line injectable disease-modifying therapies in patients with relapsing multiple sclerosis. *Curr Med Res Opin, 35*(5), 767-776. doi:10.1080/03007995.2019.1565818

Freedman, M. S., Selchen, D., Arnold, D. L., Prat, A., Banwell, B., Yeung, M., . . . Lapierre, Y. (2013). Treatment optimization in MS: Canadian MS Working Group updated recommendations. *Can J Neurol Sci, 40*(3), 307-323. doi:10.1017/s0317167100014244

Freedman, M. S., Selchen, D., Prat, A., & Giacomini, P. S. (2018). Managing Multiple Sclerosis: Treatment Initiation, Modification, and Sequencing. *Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques*, 1-15. doi:10.1017/cjn.2018.17

Friedman, C., Gatti, G., Elstein, A., Franz, T., Murphy, G., & Wolf, F. (2001). Are clinicians correct when they believe they are correct? Implications for medical decision support. *Stud Health Technol Inform, 84*(Pt 1), 454-458.

Galea, I., Lederer, C., Neuhaus, A., Muraro, P. A., Scalfari, A., Koch-Henriksen, N., . . . Daumer, M. (2013). A web-based tool for personalized prediction of long-term disease course in patients with multiple sclerosis. *Eur J Neurol, 20*(7), 1107-1109. doi:10.1111/ene.12016

Gani, R., Giovannoni, G., Bates, D., Kemball, B., Hughes, S., & Kerrigan, J. (2008). Cost-effectiveness analyses of natalizumab (Tysabri) compared with other disease-modifying therapies for people with highly active relapsing-remitting multiple sclerosis in the UK. *Pharmacoeconomics, 26*(7), 617-627. doi:10.2165/00019053-200826070-00008

Garcia Merino, A., Ramon Ara Callizo, J., Fernandez Fernandez, O., Landete Pascual, L., Moral Torres, E., & Rodriguez-Antiguedad Zarrantz, A. (2017). Consensus statement on the treatment of multiple sclerosis by the Spanish Society of Neurology in 2016. *Neurologia, 32*(2), 113-119. doi:10.1016/j.nrl.2016.02.026

Geng, J. J., Blumenfeld, Z., Tyson, T. L., & Minzenberg, M. J. (2015). Pupil diameter reflects uncertainty in attentional selection during visual search. *Front Hum Neurosci, 9*, 435. doi:10.3389/fnhum.2015.00435

Gerrity, M., White, K., DeVellis, R., & Dittus, R. (1995). Physicians' Reactions to Uncertainty: Refining the constructs and scales. *Motivation and Emotion, 19*(3), 175-191. doi:10.1007/BF02250510

Gerrity, M. S., DeVellis, R. F., & Earp, J. A. (1990). Physicians' reactions to uncertainty in patient care. A new measure and new insights. *Med Care, 28*(8), 724-736. doi:10.1097/00005650-199008000-00005

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review, 103*(3), 592-596. doi:10.1037/0033-295X.103.3.592

Giovannoni, G. (2018). Disease-modifying treatments for early and advanced multiple sclerosis: a new treatment paradigm. *Curr Opin Neurol, 31*(3), 233-243. doi:10.1097/WCO.0000000000000561

Giovannoni, G., Turner, B., Gnanapavan, S., Offiah, C., Schmierer, K., & Marta, M. (2015). Is it time to target no evident disease activity (NEDA) in multiple sclerosis? *Mult Scler Relat Disord, 4*(4), 329-333. doi:10.1016/j.msard.2015.04.006

Glimcher, P., & Fehr, E. (2014). *Neuroeconomics : Decision Making and the Brain* (Second Edition ed.). San Diego: Academic Press.

Gohlmann, S., Schmidt, C. M., & Tauchmann, H. (2010). Smoking initiation in Germany: the role of intergenerational transmission. *Health Econ, 19*(2), 227-242. doi:10.1002/hec.1470

Gongora-Ortega, J., Segovia-Bernal, Y., Valdivia-Martinez Jde, J., Galaviz-DeAnda, J. M., & Prado-Aguilar, C. A. (2012). Educational interventions to improve the effectiveness in clinical competence of general practitioners: problem-based versus critical reading-based learning. *BMC Med Educ, 12*, 53. doi:10.1186/1472-6920-12-53

Graber, M. L. (2013). The incidence of diagnostic error in medicine. *BMJ Qual Saf, 22 Suppl 2*, ii21-ii27. doi:10.1136/bmjqs-2012-001615

Gross, C. P., Vogel, E. W., Dhond, A. J., Marple, C. B., Edwards, R. A., Hauch, O., . . . Ezekowitz, M. (2003). Factors influencing physicians' reported use of anticoagulation

therapy in nonvalvular atrial fibrillation: a cross-sectional survey. *Clin Ther, 25*(6), 1750-1764.

Gupta, S., Goren, A., Phillips, A. L., Dangond, F., & Stewart, M. (2014). Self-reported severity among patients with multiple sclerosis in the U.S. and its association with health outcomes. *Mult Scler Relat Disord, 3*(1), 78-88. doi:10.1016/j.msard.2013.06.002

Gurmankin, A. D., Baron, J., Hershey, J. C., & Ubel, P. A. (2002). The role of physicians' recommendations in medical treatment decisions. *Med Decis Making, 22*(3), 262-271. doi:10.1177/0272989X0202200314

Hajek, A., Bock, J. O., & Konig, H. H. (2017). The role of personality in health care use: Results of a population-based longitudinal study in Germany. *PLoS One, 12*(7), e0181716. doi:10.1371/journal.pone.0181716

Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science (New York, N.Y.), 324*(5927), 646-648. doi:10.1126/science.1168450

Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J Neurosci, 28*(22), 5623-5630. doi:10.1523/JNEUROSCI.1309-08.2008

Harrison, M., Milbers, K., Hudson, M., & Bansback, N. (2017). Do patients and health care providers have discordant preferences about which aspects of treatments matter most? Evidence from a systematic review of discrete choice experiments. *BMJ Open, 7*(5), e014719. doi:10.1136/bmjopen-2016-014719

Hobart, J., Bowen, A., Pepper, G., Crofts, H., Eberhard, L., Berger, T., . . . Giovannoni, G. (2019). International consensus on quality standards for brain health-focused care in multiple sclerosis. *Mult Scler, 25*(13), 1809-1818. doi:10.1177/1352458518809326

Hoffmann, T., Jansen, J., & Glasziou, P. (2018). The importance and challenges of shared decision making in older people with multimorbidity. *PLoS medicine, 15*(3), e1002530-e1002530. doi:10.1371/journal.pmed.1002530

Hostalek, U., Gwilt, M., & Hildemann, S. (2015). Therapeutic Use of Metformin in Prediabetes and Diabetes Prevention. *Drugs, 75*(10), 1071-1094. doi:10.1007/s40265-015-0416-8

Hozo, I., Tsalatsanis, A., & Djulbegovic, B. (2018). Expected utility versus expected regret theory versions of decision curve analysis do generate different results when treatment effects are taken into account. *J Eval Clin Pract, 24*(1), 65-71. doi:10.1111/jep.12676

Iannello, P., Mottini, A., Tirelli, S., Riva, S., & Antonietti, A. (2017). Ambiguity and uncertainty tolerance, need for cognition, and their association with stress. A study among Italian practicing physicians. *Med Educ Online, 22*(1), 1270009. doi:10.1080/10872981.2016.1270009

iMotions. (2016). *Facial Expression Analysis*. Retrieved from Boston, MA 02111: https://imotions.com/wp-content/uploads/Guides/iMotions_Guide_FacialExpressions_2016.pdf

Ioannidis, J. P., & Lau, J. (2001). Evidence on interventions to reduce medical errors: an overview and recommendations for future research. *J Gen Intern Med, 16*(5), 325-334.

Jongen, P. J., Heerings, M., Lemmens, W. A., Donders, R., van der Zande, A., van Noort, E., & Kool, A. (2015). A prospective web-based patient-centred interactive study of long-term disabilities, disabilities perception and health-related quality of life in patients with multiple sclerosis in The Netherlands: the Dutch Multiple Sclerosis Study protocol. *BMC Neurol, 15*, 128. doi:10.1186/s12883-015-0379-0

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron, 89*(1), 221-234. doi:10.1016/j.neuron.2015.11.028

Keren, G., & Schul, Y. (2009). Two Is Not Always Better Than One: A Critical Evaluation of Two-System Theories. *Perspect Psychol Sci, 4*(6), 533-550. doi:10.1111/j.1745-6924.2009.01164.x

Kerr, E. A., Zikmund-Fisher, B. J., Klamerus, M. L., Subramanian, U., Hogan, M. M., & Hofer, T. P. (2008). The role of clinical uncertainty in treatment decisions for diabetic patients with uncontrolled blood pressure. *Ann Intern Med, 148*(10), 717-727.

Khunti, K., Nikolajsen, A., Thorsted, B. L., Andersen, M., Davies, M. J., & Paul, S. K. (2016). Clinical inertia with regard to intensifying therapy in people with type 2 diabetes treated with basal insulin. *Diabetes Obes Metab, 18*(4), 401-409. doi:10.1111/dom.12626

Klucken, T., Schweckendiek, J., Koppe, G., Merz, C. J., Kagerer, S., Walter, B., . . . Stark, R. (2012). Neural correlates of disgust- and fear-conditioned responses. *Neuroscience, 201*, 209-218. doi:https://doi.org/10.1016/j.neuroscience.2011.11.007

Kobelt, G., Berg, J., Lindgren, P., Fredrikson, S., & Jonsson, B. (2006). Costs and quality of life of patients with multiple sclerosis in Europe. *J Neurol Neurosurg Psychiatry, 77*(8), 918-926. doi:10.1136/jnnp.2006.090365

Kostopoulou, O., Russo, J. E., Keenan, G., Delaney, B. C., & Douiri, A. (2012). Information distortion in physicians' diagnostic judgments. *Med Decis Making, 32*(6), 831-839. doi:10.1177/0272989X12447241

Lambert, C. T., Bumgarner, J. M., & Tarakji, K. G. (2019). Atrial Fibrillation Detection With Wearable Devices. *JAMA, 321*(23), 2367-2368. doi:10.1001/jama.2019.4538

Lavin, C., San Martin, R., & Rosales Jubal, E. (2013). Pupil dilation signals uncertainty and surprise in a learning gambling task. *Front Behav Neurosci, 7*, 218. doi:10.3389/fnbeh.2013.00218

Lebeau, J. P., Cadwallader, J. S., Aubin-Auger, I., Mercier, A., Pasquet, T., Rusch, E., . . . Vermeire, E. (2014). The concept and definition of therapeutic inertia in hypertension in primary care: a qualitative systematic review. *BMC Fam Pract, 15*, 130. doi:10.1186/1471-2296-15-130

Lebeau, J. P., Cadwallader, J. S., Vaillant-Roussel, H., Pouchain, D., Yaouanc, V., Aubin-Auger, I., . . . Hendrickx, K. (2016). General practitioners' justifications for therapeutic inertia in cardiovascular prevention: an empirically grounded typology. *BMJ Open, 6*(5), e010639. doi:10.1136/bmjopen-2015-010639

Légaré, F., & Witteman, H. O. (2013). Shared Decision Making: Examining Key Elements And Barriers To Adoption Into Routine Clinical Practice. *Health Affairs, 32*(2), 276-284. doi:10.1377/hlthaff.2012.1078

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and Decision Making. *Annual Review of Psychology, 66*(1), 799-823. doi:10.1146/annurev-psych-010213-115043

Levy, I., Snell, J., Nelson, A. J., Rustichini, A., & Glimcher, P. W. (2010). Neural representation of subjective value under risk and ambiguity. *J Neurophysiol, 103*(2), 1036-1047. doi:10.1152/jn.00853.2009

Lewis, J. R., & Sauro, J. (2009, 2009//). *The Factor Structure of the System Usability Scale.* Paper presented at the Human Centered Design, Berlin, Heidelberg.

Lown, M., Brown, M., Brown, C., Yue, A. M., Shah, B. N., Corbett, S. J., . . . Little, P. (2020). Machine learning detection of Atrial Fibrillation using wearable technology. *PLoS One, 15*(1), e0227401. doi:10.1371/journal.pone.0227401

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation Analysis. *Annual Review of Psychology, 58*(1), 593-614. doi:10.1146/annurev.psych.58.110405.085542

Mahad, D. H., Trapp, B. D., & Lassmann, H. (2015). Pathological mechanisms in progressive multiple sclerosis. *Lancet Neurol, 14*(2), 183-193. doi:10.1016/S1474-4422(14)70256-X

Maier, M. E., Ernst, B., & Steinhauser, M. (2019). Error-related pupil dilation is sensitive to the evaluation of different error types. *Biol Psychol, 141*, 25-34. doi:10.1016/j.biopsycho.2018.12.013

Makary, M. A., & Daniel, M. (2016). Medical error—the third leading cause of death in the US. *BMJ, 353*, i2139. doi:10.1136/bmj.i2139

Mamede, S., Schmidt, H. G., Rikers, R. M., Custers, E. J., Splinter, T. A., & van Saase, J. L. (2010). Conscious thought beats deliberation without attention in diagnostic decision-making: at least when you are an expert. *Psychol Res, 74*(6), 586-592. doi:10.1007/s00426-010-0281-8

Mamede, S., van Gog, T., van den Berge, K., Rikers, R. M., van Saase, J. L., van Guldener, C., & Schmidt, H. G. (2010). Effect of availability bias and reflective reasoning on diagnostic accuracy among internal medicine residents. *JAMA, 304*(11), 1198-1203. doi:10.1001/jama.2010.1276

Mamede, S., van Gog, T., van den Berge, K., van Saase, J. L., & Schmidt, H. G. (2014). Why do doctors make mistakes? A study of the role of salient distracting clinical features. *Acad Med, 89*(1), 114-120. doi:10.1097/ACM.0000000000000077

Manski-Nankervis, J. A., Furler, J., O'Neal, D., Ginnivan, L., Thuraisingam, S., & Blackberry, I. (2017). Overcoming clinical inertia in insulin initiation in primary care for patients with type 2 diabetes: 24-month follow-up of the Stepping Up cluster randomised controlled trial. *Prim Care Diabetes, 11*(5), 474-481. doi:10.1016/j.pcd.2017.06.005

Mata-Cases, M., Franch-Nadal, J., Real, J., Gratacos, M., Lopez-Simarro, F., Khunti, K., & Mauricio, D. (2018). Therapeutic inertia in patients treated with two or more antidiabetics in primary care: Factors predicting intensification of treatment. *Diabetes Obes Metab, 20*(1), 103-112. doi:10.1111/dom.13045

Mataix-Cols, D., An, S. K., Lawrence, N. S., Caseras, X., Speckens, A., Giampietro, V., . . . Phillips, M. L. (2008). Individual differences in disgust sensitivity modulate neural responses to aversive/disgusting stimuli. *European Journal of Neuroscience, 27*(11), 3050-3058. doi:doi:10.1111/j.1460-9568.2008.06311.x

Mathot, S. (2018). Pupillometry: Psychology, Physiology, and Function. *J Cogn, 1*(1), 16. doi:10.5334/joc.18

Matute-Blanch, C., Montalban, X., & Comabella, M. (2017). Multiple sclerosis, and other demyelinating and autoimmune inflammatory diseases of the central nervous system. *Handb Clin Neurol, 146*, 67-84. doi:10.1016/B978-0-12-804279-3.00005-8

McCrone, P., Heslin, M., Knapp, M., Bull, P., & Thompson, A. (2008). Multiple sclerosis in the UK: service use, costs, quality of life and disability. *Pharmacoeconomics, 26*(10), 847-860. doi:10.2165/00019053-200826100-00005

McDuff, D., Kodra, E., Kaliouby, R. E., & LaFrance, M. (2017). A large-scale analysis of sex differences in facial expressions. *PLoS One, 12*(4), e0173942. doi:10.1371/journal.pone.0173942

Meinema, J. G., Buwalda, N., van Etten-Jamaludin, F. S., Visser, M. R. M., & van Dijk, N. (2019). Intervention Descriptions in Medical Education: What Can Be Improved? A Systematic Review and Checklist. *Acad Med, 94*(2), 281-290. doi:10.1097/ACM.0000000000002428

Meyer, A. N., Payne, V. L., Meeks, D. W., Rao, R., & Singh, H. (2013). Physicians' diagnostic accuracy, confidence, and resource requests: a vignette study. *JAMA Intern Med, 173*(21), 1952-1958. doi:10.1001/jamainternmed.2013.10081

Milman, T., Joundi, R. A., Alotaibi, N. M., & Saposnik, G. (2018). Clinical inertia in the pharmacological management of hypertension: A systematic review and meta-analysis. *Medicine (Baltimore), 97*(25), e11121. doi:10.1097/MD.0000000000011121

Mohan, A. V., & Phillips, L. S. (2011). Clinical inertia and uncertainty in medicine. *JAMA, 306*(4), 383; author reply 383-384. doi:10.1001/jama.2011.1044

Monrouxe, L. V., Grundy, L., Mann, M., John, Z., Panagoulas, E., Bullock, A., & Mattick, K. (2017). How prepared are UK medical graduates for practice? A rapid review of the literature 2009-2014. *BMJ Open, 7*(1), e013656. doi:10.1136/bmjopen-2016-013656

Montalban, X., Gold, R., Thompson, A. J., Otero-Romero, S., Amato, M. P., Chandraratna, D., . . . Zipp, F. (2018). ECTRIMS/EAN Guideline on the pharmacological treatment of people with multiple sclerosis. *Mult Scler, 24*(2), 96-120. doi:10.1177/1352458517751049

Muhammad Sayem, A. S., Hon Teay, S., Shahariar, H., Fink, P. L., & Albarbar, A. (2020). Review on Smart Electro-Clothing Systems (SeCSs). *Sensors (Basel), 20*(3). doi:10.3390/s20030587

Mukherjee, K. (2010). A dual system model of preferences under risk. *Psychological Review, 117*(1), 243.

Murphy, M. S., & Baker, M. (2014). When used appropriately, NICE's traffic light system for febrile children helps clinicians make safe and appropriate decisions. *BMJ, 348*, g2518. doi:10.1136/bmj.g2518

Murphy, P. R., O'Connell, R. G., O'Sullivan, M., Robertson, I. H., & Balsters, J. H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus. *Human brain mapping, 35*(8), 4140-4154.

Murphy, P. R., Vandekerckhove, J., & Nieuwenhuis, S. (2014). Pupil-linked arousal determines variability in perceptual decision making. *PLoS Comput Biol, 10*(9), e1003854. doi:10.1371/journal.pcbi.1003854

Ness, N. H., Haase, R., Kern, R., Schriefer, D., Ettle, B., Cornelissen, C., . . . Ziemssen, T. (2020). The Multiple Sclerosis Health Resource Utilization Survey (MS-HRS): Development and Validation Study. *J Med Internet Res, 22*(3), e17921. doi:10.2196/17921

Ness, N. H., Schriefer, D., Haase, R., Ettle, B., Cornelissen, C., & Ziemssen, T. (2020). Differentiating societal costs of disability worsening in multiple sclerosis. *J Neurol, 267*(4), 1035-1042. doi:10.1007/s00415-019-09676-4

Noyes, K., & Weinstock-Guttman, B. (2013). Impact of diagnosis and early treatment on the course of multiple sclerosis. *Am J Manag Care, 19*(17 Suppl), s321-331.

O'Connor, P. J., Sperl-Hillen, J. A. M., Johnson, P. E., Rush, W. A., & Biltz, G. (2005). Clinical Inertia and Outpatient Medical Errors. In K. Henriksen, J. B. Battles, E. S. Marks, & D. I. Lewin (Eds.), *Advances in Patient Safety: From Research to Implementation (Volume 2: Concepts and Methodology)*. Rockville (MD).

Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion. *Trends in cognitive sciences, 9*(5), 242-249. doi:10.1016/j.tics.2005.03.010

OECD. *Health at a Glance 2013*: OECD Publishing.

Ogura, M., & Harada-Shiba, M. (2016). Clinical Inertia in the Management of Hypercholesterolemia: What Clinicians Need to do. *J Atheroscler Thromb, 23*(5), 552-553. doi:10.5551/jat.ED037

Oh, J., Vidal-Jordana, A., & Montalban, X. (2018). Multiple sclerosis: clinical aspects. *Current Opinion in Neurology, 31*(6), 752-759. doi:10.1097/wco.0000000000000622

Okonofua, E. C., Simpson, K. N., Jesri, A., Rehman, S. U., Durkalski, V. L., & Egan, B. M. (2006). Therapeutic inertia is an impediment to achieving the Healthy People 2010 blood pressure control goals. *Hypertension, 47*(3), 345-351. doi:10.1161/01.HYP.0000200702.76436.4b

Ontaneda, D., Cohn, S., & Fox, R. J. Risk stratification and mitigation in multiple sclerosis. *Multiple Sclerosis and Related Disorders, 3*(5), 639-649. doi:10.1016/j.msard.2014.05.003

Ontaneda, D., Fox, R. J., & Chataway, J. (2015). Clinical trials in progressive multiple sclerosis: lessons learned and future perspectives. *Lancet Neurol, 14*(2), 208-223. doi:10.1016/S1474-4422(14)70264-9

Ontaneda, D., Tallantyre, E., Kalincik, T., Planchon, S. M., & Evangelou, N. (2019). Early highly effective versus escalation treatment approaches in relapsing multiple sclerosis. *Lancet Neurol, 18*(10), 973-980. doi:10.1016/S1474-4422(19)30151-6

Orriols, L., Delorme, B., Gadegbeku, B., Tricotel, A., Contrand, B., Laumon, B., . . . group, C. r. (2010). Prescription medicines and the risk of road traffic crashes: a French registry-based study. *PLoS Med, 7*(11), e1000366. doi:10.1371/journal.pmed.1000366

Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychon Bull Rev, 11*(6), 988-1010. doi:10.3758/bf03196730

Otto, B., Misra, S., Prasad, A., & McRae, K. (2014). Functional overlap of top-down emotion regulation and generation: An fMRI study identifying common neural substrates between cognitive reappraisal and cognitively generated emotions. *Cognitive, Affective, & Behavioral Neuroscience, 14*(3), 923-938. doi:10.3758/s13415-013-0240-0

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature, 441*(7090), 223-226. doi:10.1038/nature04676

Paulus, M. P., & Yu, A. J. (2012). Emotion and decision-making: affect-driven belief systems in anxiety and depression. *Trends in cognitive sciences, 16*(9), 476-483. doi:10.1016/j.tics.2012.07.009

Peake, J. M., Kerr, G., & Sullivan, J. P. (2018). A Critical Review of Consumer Wearables, Mobile Applications, and Equipment for Providing Biofeedback, Monitoring Stress, and Sleep in Physically Active Populations. *Frontiers in Physiology, 9*(743). doi:10.3389/fphys.2018.00743

Perneger, T. V., & Agoritsas, T. (2011). Doctors and patients' susceptibility to framing bias: a randomized trial. *J Gen Intern Med, 26*(12), 1411-1417. doi:10.1007/s11606-011-1810-x

Phelps, E. A., Lempert, K. M., & Sokol-Hessner, P. (2014). Emotion and Decision Making: Multiple Modulatory Neural Circuits. *Annual Review of Neuroscience, 37*(1), 263-287. doi:10.1146/annurev-neuro-071013-014119

Phillips, A. C., Lewis, L. K., McEvoy, M. P., Galipeau, J., Glasziou, P., Moher, D., . . . Williams, M. T. (2016). Development and validation of the guideline for reporting evidence-based practice educational interventions and teaching (GREET). *BMC Med Educ, 16*(1), 237. doi:10.1186/s12909-016-0759-1

Phillips, L. S., Branch, W. T., Cook, C. B., Doyle, J. P., El-Kebbi, I. M., Gallina, D. L., . . . Barnes, C. S. (2001). Clinical inertia. *Ann Intern Med, 135*(9), 825-834.

Plassmann, H., O'Doherty, J., Shiv, B., & Rangel, A. (2008). Marketing actions can modulate neural representations of experienced pleasantness. *Proc Natl Acad Sci U S A, 105*(3), 1050-1054. doi:10.1073/pnas.0706929105

Platt, M. L., & Huettel, S. A. (2008). Risky business: the neuroeconomics of decision making under uncertainty. *Nature neuroscience, 11*(4), 398-403. doi:10.1038/nn2062

Prakash, S., Sladek, R. M., & Schuwirth, L. (2019). Interventions to improve diagnostic decision making: A systematic review and meta-analysis on reflective strategies. *Med Teach, 41*(5), 517-524. doi:10.1080/0142159X.2018.1497786

Preuschoff, K., t Hart, B. M., & Einhauser, W. (2011). Pupil Dilation Signals Surprise: Evidence for Noradrenaline's Role in Decision Making. *Front Neurosci, 5*, 115. doi:10.3389/fnins.2011.00115

Privitera, C. M., Renninger, L. W., Carney, T., Klein, S., & Aguilar, M. (2010). Pupil dilation during visual target detection. *Journal of Vision, 10*(10), 3-3. doi:10.1167/10.10.3

Prosperini, L., Gianni, C., Leonardi, L., De Giglio, L., Borriello, G., Galgani, S., . . . Gasperini, C. (2012). Escalation to natalizumab or switching among immunomodulators in relapsing multiple sclerosis. *Mult Scler, 18*(1), 64-71. doi:10.1177/1352458511417481

Prosperini, L., Mancinelli, C. R., De Giglio, L., De Angelis, F., Barletta, V., & Pozzilli, C. (2014). Interferon beta failure predicted by EMA criteria or isolated MRI activity in multiple sclerosis. *Mult Scler, 20*(5), 566-576. doi:10.1177/1352458513502399

Rae-Grant, A., Day, G. S., Marrie, R. A., Rabinstein, A., Cree, B. A. C., Gronseth, G. S., . . . Pringsheim, T. (2018). Practice guideline recommendations summary: Disease-modifying therapies for adults with multiple sclerosis: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. *Neurology, 90*(17), 777-788. doi:10.1212/WNL.0000000000005347

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature reviews. Neuroscience, 9*(7), 545-556. doi:10.1038/nrn2357

Reach, G. (2014). Clinical inertia, uncertainty and individualized guidelines. *Diabetes Metab, 40*(4), 241-245. doi:10.1016/j.diabet.2013.12.009

Redelmeier, D. A., & Shafir, E. (1995). Medical decision making in situations that offer multiple alternatives. *JAMA, 273*(4), 302-305.

Reyna, V. F., & Lloyd, F. J. (2006). Physician decision making and cardiac risk: effects of knowledge, risk perception, risk tolerance, and fuzzy processing. *J Exp Psychol Appl, 12*(3), 179-195. doi:10.1037/1076-898X.12.3.179

Richards, R. G., Sampson, F. C., Beard, S. M., & Tappenden, P. (2002). A review of the natural history and epidemiology of multiple sclerosis: implications for resource allocation and health economic models. *Health Technol Assess, 6*(10), 1-73. doi:10.3310/hta6100

Ritov, I., & Baron, J. (1992). Status-quo and omission biases. *Journal of risk and uncertainty, 5*(1), 49-61.

Rotstein, D., & Montalban, X. (2019). Reaching an evidence-based prognosis for personalized treatment of multiple sclerosis. *Nat Rev Neurol, 15*(5), 287-300. doi:10.1038/s41582-019-0170-8

Saposnik, G., Cote, R., Mamdani, M., Raptis, S., Thorpe, K. E., Fang, J., . . . Goldstein, L. B. (2013). JURaSSiC: accuracy of clinician vs risk score prediction of ischemic stroke outcomes. *Neurology, 81*(5), 448-455. doi:10.1212/WNL.0b013e31829d874e

Saposnik, G., Mamdani, M., Montalban, X., Terzaghi, M., Silva, B., Saladino, M. L., . . . Caceres, F. (2019). Traffic Lights Intervention Reduces Therapeutic Inertia: A Randomized Controlled Trial in Multiple Sclerosis Care. *MDM Policy Pract, 4*(1), 2381468319855642. doi:10.1177/2381468319855642

Saposnik, G., Maurino, J., Sempere, A. P., Terzaghi, M. A., Ruff, C. C., Mamdani, M., . . . Montalban, X. (2017). Overcoming Therapeutic Inertia in Multiple Sclerosis Care: A Pilot Randomized Trial Applying the Traffic Light System in Medical Education. *Front Neurol, 8*, 430. doi:10.3389/fneur.2017.00430

Saposnik, G., & Montalban, X. (2018). Therapeutic Inertia in the New Landscape of Multiple Sclerosis Care. *Front Neurol, 9*, 174. doi:10.3389/fneur.2018.00174

Saposnik, G., Montalban, X., Selchen, D., Terzaghi, M. A., Bakdache, F., Montoya, A., . . . Oh, J. (2018). Therapeutic Inertia in Multiple Sclerosis Care: A Study of Canadian Neurologists. *Front Neurol, 9*, 781. doi:10.3389/fneur.2018.00781

Saposnik, G., Oh, J., Terzaghi, M. A., Kostyrko, P., Bakdache, F., Montoya, A., . . . Tobler, P. N. (2019). Emotional expressions associated with therapeutic inertia in multiple sclerosis care. *Mult Scler Relat Disord, 34*, 17-28. doi:10.1016/j.msard.2019.05.029

Saposnik, G., Redelmeier, D., Ruff, C. C., & Tobler, P. N. (2016). Cognitive biases associated with medical decisions: a systematic review. *BMC Med Inform Decis Mak, 16*(1), 138. doi:10.1186/s12911-016-0377-1

Saposnik, G., Sempere, A. P., Prefasi, D., Selchen, D., Ruff, C. C., Maurino, J., & Tobler, P. N. (2017). Decision-making in Multiple Sclerosis: The Role of Aversion to Ambiguity for Therapeutic Inertia among Neurologists (DIScUTIR MS). *Front Neurol, 8*, 65. doi:10.3389/fneur.2017.00065

Saposnik, G., Sempere, A. P., Raptis, R., Prefasi, D., Selchen, D., & Maurino, J. (2016a). Decision making under uncertainty, therapeutic inertia, and physicians' risk preferences in the management of multiple sclerosis (DIScUTIR MS). *BMC Neurol, 16*, 58. doi:10.1186/s12883-016-0577-4

Saposnik, G., Sempere, A. P., Raptis, R., Prefasi, D., Selchen, D., & Maurino, J. (2016b). Decision making under uncertainty, therapeutic inertia, and physicians' risk preferences in the management of multiple sclerosis (DIScUTIR MS). *BMC Neurol, 16*(1), 58. doi:10.1186/s12883-016-0577-4

Saposnik, G., Tobler, P. N., Caceres, F., Terzaghi, M. A., Ruff, C., Maurino, J., . . . Mamdani, M. (2018). Usability of an Educational Intervention to Overcome Therapeutic Inertia in Multiple Sclerosis Care. *Front Neurol, 9*, 522. doi:10.3389/fneur.2018.00522

Schwartz, A. (2011). Medical decision making and medical education: challenges and opportunities. *Perspect Biol Med, 54*(1), 68-74. doi:10.1353/pbm.2011.0010

Schwartz, A. (2013). A consensus on concussion: While there is agreement on guidelines, there are gaps in awareness and education. *Ann Neurol, 74*(4), A26-27. doi:10.1002/ana.24031

Semaan, S., Dewland, T. A., Tison, G. H., Nah, G., Vittinghoff, E., Pletcher, M. J., . . . Marcus, G. M. (2020). Physical activity and atrial fibrillation: Data from wearable fitness trackers. *Heart Rhythm, 17*(5 Pt B), 842-846. doi:10.1016/j.hrthm.2020.02.013

Shevchenko, Y., von Helversen, B., & Scheibehenne, B. (2014). Change and status quo in decisions with defaults: The effect of incidental emotions depends on the type of default. *Judgment and Decision Making, 9*(3), 287-296.

Shi, H., Zhao, H., Liu, Y., Gao, W., & Dou, S. C. (2019). Systematic Analysis of a Military Wearable Device Based on a Multi-Level Fusion Framework: Research Directions. *Sensors (Basel), 19*(12). doi:10.3390/s19122651

Sibbald, M., de Bruin, A. B., & van Merrienboer, J. J. (2013). Checklists improve experts' diagnostic decisions. *Med Educ, 47*(3), 301-308. doi:10.1111/medu.12080

Sokol-Hessner, P., Camerer, C. F., & Phelps, E. A. (2013). Emotion regulation reduces loss aversion and decreases amygdala responses to losses. *Social Cognitive and Affective Neuroscience, 8*(3), 341-350. doi:10.1093/scan/nss002

Sonnenberg, L., Gelsomin, E., Levy, D. E., Riis, J., Barraclough, S., & Thorndike, A. N. (2013). A traffic light food labeling intervention increases consumer awareness of health and healthy choices at the point-of-purchase. *Preventive Medicine, 57*(4), 253-257. doi:http://dx.doi.org/10.1016/j.ypmed.2013.07.001

Sormani, M. P., & Bruzzi, P. (2015). Can we measure long-term treatment effects in multiple sclerosis? *Nat Rev Neurol, 11*(3), 176-182. doi:10.1038/nrneurol.2014.237

Sormani, M. P., Gasperini, C., Romeo, M., Rio, J., Calabrese, M., Cocco, E., . . . group, M. s. (2016). Assessing response to interferon-beta in a multicenter dataset of patients with MS. *Neurology, 87*(2), 134-140. doi:10.1212/WNL.0000000000002830

Sormani, M. P., Rio, J., Tintore, M., Signori, A., Li, D., Cornelisse, P., . . . De Stefano, N. (2013). Scoring treatment response in patients with relapsing multiple sclerosis. *Mult Scler, 19*(5), 605-612. doi:10.1177/1352458512460605

Sorum, P. C., Shim, J., Chasseigne, G., Bonnin-Scaon, S., Cogneau, J., & Mullet, E. (2003). Why do primary care physicians in the United States and France order prostate-specific antigen tests for asymptomatic patients? *Med Decis Making, 23*(4), 301-313.

Stacey, D., Legare, F., Lewis, K., Barry, M. J., Bennett, C. L., Eden, K. B., . . . Trevena, L. (2017). Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev, 4*, CD001431. doi:10.1002/14651858.CD001431.pub5

Stangel, M., Penner, I. K., Kallmann, B., Lukas, C., Kieseier, B., & Gold, R. (2014). Multiple Sclerosis Decision Model (MSDM): A Multifactorial Model To Monitor Treatment Response And Disease Course In Relapsing Remitting Multiple Sclerosis (P3.133). *Neurology, 82*(10 Supplement), P3.133.

Stangierski, A., Warmuz-Stangierska, I., Ruchala, M., Zdanowska, J., Glowacka, M. D., Sowinski, J., & Ruchala, P. (2012). Medical errors - not only patients' problem. *Arch Med Sci, 8*(3), 569-574. doi:10.5114/aoms.2012.29413

Stiegler, M. P., & Ruskin, K. J. (2012). Decision-making and safety in anesthesiology. *Curr Opin Anaesthesiol, 25*(6), 724-729. doi:10.1097/ACO.0b013e328359307a

Stöckli, S., Schulte-Mecklenbeck, M., Borer, S., & Samson, A. C. (2018). Facial expression analysis with AFFDEX and FACET: A validation study. *Behavior Research Methods, 50*(4), 1446-1460. doi:10.3758/s13428-017-0996-1

Stolper, E., Van de Wiel, M., Van Royen, P., Van Bokhoven, M., Van der Weijden, T., & Dinant, G. J. (2011). Gut feelings as a third track in general practitioners' diagnostic reasoning. *J Gen Intern Med, 26*(2), 197-203. doi:10.1007/s11606-010-1524-5

Stripe, S. C., Best, L. G., Cole-Harding, S., Fifield, B., & Talebdoost, F. (2006). Aviation model cognitive risk factors applied to medical malpractice cases. *J Am Board Fam Med, 19*(6), 627-632.

Summerfield, C., & Tsetsos, K. (2015). Do humans make good decisions? *Trends in cognitive sciences, 19*(1), 27-34. doi:10.1016/j.tics.2014.11.005

Thaler, R. H., & Sunstein, C. R. (2008). Nudge: improving decisions about health. *Wealth, and Happiness, 6*.

Tramacere, I., Del Giovane, C., Salanti, G., D'Amico, R., & Filippini, G. (2015). Immunomodulators and immunosuppressants for relapsing-remitting multiple sclerosis: a network meta-analysis. *Cochrane Database Syst Rev*(9), CD011381. doi:10.1002/14651858.CD011381.pub2

Tsalatsanis, A., Hozo, I., Kumar, A., & Djulbegovic, B. (2015). Dual Processing Model for Medical Decision-Making: An Extension to Diagnostic Testing. *PLoS One, 10*(8), e0134800. doi:10.1371/journal.pone.0134800

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science (New York, N.Y.), 185*(4157), 1124-1131. doi:10.1126/science.185.4157.1124

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science (New York, N.Y.), 211*(4481), 453-458. doi:10.1126/science.7455683

Urai, A. E., Braun, A., & Donner, T. H. (2017). Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. *Nat Commun, 8*, 14637. doi:10.1038/ncomms14637

van den Berge, K., & Mamede, S. (2013). Cognitive diagnostic error in internal medicine. *Eur J Intern Med, 24*(6), 525-529. doi:10.1016/j.ejim.2013.03.006

van Hooff, J. C., van Buuringen, M., El M'rabet, I., de Gier, M., & van Zalingen, L. (2014). Disgust-specific modulation of early attention processes. *Acta Psychologica, 152*, 149-157. doi:https://doi.org/10.1016/j.actpsy.2014.08.009

VanderWeele, T. J. (2016). Mediation Analysis: A Practitioner's Guide. *Annual Review of Public Health, 37*(1), 17-32. doi:10.1146/annurev-publhealth-032315-021402

Vidal-Jordana, A., & Montalban, X. (2017). Multiple Sclerosis: Epidemiologic, Clinical, and Therapeutic Aspects. *Neuroimaging Clin N Am, 27*(2), 195-204. doi:10.1016/j.nic.2016.12.001

Vilares, I., & Kording, K. (2011). Bayesian models: the structure of the world, uncertainty, behavior, and the brain. *Annals of the New York Academy of Sciences, 1224*(1), 22-39. doi:10.1111/j.1749-6632.2011.05965.x

Wang, C.-A., Baird, T., Huang, J., Coutinho, J. D., Brien, D. C., & Munoz, D. P. (2018). Arousal Effects on Pupil Size, Heart Rate, and Skin Conductance in an Emotional Face Task. *Frontiers in Neurology, 9*(1029). doi:10.3389/fneur.2018.01029

Wang, C. A., & Munoz, D. P. (2015). A circuit for pupil orienting responses: implications for cognitive modulation of pupil size. *Curr Opin Neurobiol, 33*, 134-140. doi:10.1016/j.conb.2015.03.018

Wells G, Shea B, O'Connell D, Peterson J, Welch V, Losos M, & P, T. (2013). The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Retrieved from In http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

Wessel, J. R., Dolan, K. A., & Hollingworth, A. (2018). A blunted phasic autonomic response to errors indexes age-related deficits in error awareness. *Neurobiol Aging, 71*, 13-20. doi:10.1016/j.neurobiolaging.2018.06.019

Yee, L. M., Liu, L. Y., & Grobman, W. A. (2014). The relationship between obstetricians' cognitive and affective traits and their patients' delivery outcomes. *Am J Obstet Gynecol, 211*(6), 692 e691-696. doi:10.1016/j.ajog.2014.06.003

Zeltser, M. V., & Nash, D. B. (2010). Approaching the evidence basis for aviation-derived teamwork training in medicine. *Am J Med Qual, 25*(1), 13-23. doi:10.1177/1062860609345664

Zhang, L., Wang, K., Zhu, C., Yu, F., & Chen, X. (2015). Trait Anxiety Has Effect on Decision Making under Ambiguity but Not Decision Making under Risk. *PLoS One, 10*(5), e0127189-e0127189. doi:10.1371/journal.pone.0127189

Zhang, X., Liu, Y., Gu, Y., Wang, S., & Chen, H. (2020). Red for "Stop": "Traffic-Light" Nutrition Labels Decrease Unhealthy Food Choices by Increasing Activity and Connectivity in the Frontal Lobe. *Nutrients, 12*(1). doi:10.3390/nu12010128

Zikmund-Fisher, B. J., Couper, M. P., Singer, E., Levin, C. A., Fowler, F. J., Ziniel, S., . . . Fagerlin, A. (2010). The DECISIONS Study: A Nationwide Survey of United States Adults Regarding 9 Common Medical Decisions. *Medical Decision Making, 30*(5_suppl), 20-34. doi:10.1177/0272989x09353792

Zwaan, L., Thijs, A., Wagner, C., van der Wal, G., & Timmermans, D. R. (2012). Relating faults in diagnostic reasoning with diagnostic errors and patient harm. *Acad Med, 87*(2), 149-156. doi:10.1097/ACM.0b013e31823f71e6

## List of abbreviations

**CT** – Computerized tomography scan
**CI** – Confidence intervals
**dlPFC** – dorsolateral prefrontal cortex
**DMT** – disease-modifying therapies
**DPT** – Dual process theory
**EDSS** –Expanded Disability Status Scale
**EUT** – Expected utility theory
**FDA** - Federal Drug Agency
**GEE** – Generalized estimated equations
**fMRI** – Functional magnetic resonance imaging
**MA** – European Medical Agency
**MS** – Multiple sclerosis
**NEDA** – no evidence of disease activity
**OR** – Odds ratio
**SD** – standard deviation
**SDM** – Shared decision making
**SEM** – Standard error of the mean
**SOEP** - German Socio-Economic Panel
**SQ** – Status quo
**SUS** – System usability score
**TI** – Therapeutic inertia
**TLS** – Traffic light system
**UZH** – University of Zurich
**vmPFC** – ventromedial prefrontal cortex

# Appendices

# A. Appendix to Study 1

Saposnik et al. BMC Medical Informatics and Decision Making (2016) 16:138
DOI 10.1186/s12911-016-0377-1

BMC Medical Informatics and Decision Making

RESEARCH ARTICLE Open Access

## Cognitive biases associated with medical decisions: a systematic review

Gustavo Saposnik[1,2,3,4*], Donald Redelmeier[3], Christian C. Ruff[1†] and Philippe N. Tobler[1†]

**Abstract**

**Background:** Cognitive biases and personality traits (aversion to risk or ambiguity) may lead to diagnostic inaccuracies and medical errors resulting in mismanagement or inadequate utilization of resources. We conducted a systematic review with four objectives: 1) to identify the most common cognitive biases, 2) to evaluate the influence of cognitive biases on diagnostic accuracy or management errors, 3) to determine their impact on patient outcomes, and 4) to identify literature gaps.

**Methods:** We searched MEDLINE and the Cochrane Library databases for relevant articles on cognitive biases from 1980 to May 2015. We included studies conducted in physicians that evaluated at least one cognitive factor using case-vignettes or real scenarios and reported an associated outcome written in English. Data quality was assessed by the Newcastle-Ottawa scale. Among 114 publications, 20 studies comprising 6810 physicians met the inclusion criteria. Nineteen cognitive biases were identified.

**Results:** All studies found at least one cognitive bias or personality trait to affect physicians. Overconfidence, lower tolerance to risk, the anchoring effect, and information and availability biases were associated with diagnostic inaccuracies in 36.5 to 77 % of case-scenarios. Five out of seven (71.4 %) studies showed an association between cognitive biases and therapeutic or management errors. Of two (10 %) studies evaluating the impact of cognitive biases or personality traits on patient outcomes, only one showed that higher tolerance to ambiguity was associated with increased medical complications (9.7 % vs 6.5 %; $p = .004$). Most studies (60 %) targeted cognitive biases in diagnostic tasks, fewer focused on treatment or management (35 %) and on prognosis (10 %). Literature gaps include potentially relevant biases (e.g. aggregate bias, feedback sanction, hindsight bias) not investigated in the included studies. Moreover, only five (25 %) studies used clinical guidelines as the framework to determine diagnostic or treatment errors. Most studies ($n = 12$, 60 %) were classified as low quality.

**Conclusions:** Overconfidence, the anchoring effect, information and availability bias, and tolerance to risk may be associated with diagnostic inaccuracies or suboptimal management. More comprehensive studies are needed to determine the prevalence of cognitive biases and personality traits and their potential impact on physicians' decisions, medical errors, and patient outcomes.

**Keywords:** Decision making, Cognitive bias, Personality traits, Cognition, Physicians, Case-scenarios, Systematic review

This paper is published as Saposnik G, Redelmeier D, Ruff CC, Tobler PN, Cognitive biases associated with medical decisions: a systematic review. BMC Med Inform Decis Mak.,16(1):138, 2016.

Saposnik et al. BMC Medical Informatics and Decision Making (2016) 16:138
DOI 10.1186/s12911-016-0377-1

BMC Medical Informatics and Decision Making

RESEARCH ARTICLE Open Access

# Cognitive biases associated with medical decisions: a systematic review

Gustavo Saposnik[1,2,3,4*], Donald Redelmeier[3], Christian C. Ruff[1†] and Philippe N. Tobler[1†]

## Abstract

**Background:** Cognitive biases and personality traits (aversion to risk or ambiguity) may lead to diagnostic inaccuracies and medical errors resulting in mismanagement or inadequate utilization of resources. We conducted a systematic review with four objectives: 1) to identify the most common cognitive biases, 2) to evaluate the influence of cognitive biases on diagnostic accuracy or management errors, 3) to determine their impact on patient outcomes, and 4) to identify literature gaps.

**Methods:** We searched MEDLINE and the Cochrane Library databases for relevant articles on cognitive biases from 1980 to May 2015. We included studies conducted in physicians that evaluated at least one cognitive factor using case-vignettes or real scenarios and reported an associated outcome written in English. Data quality was assessed by the Newcastle-Ottawa scale. Among 114 publications, 20 studies comprising 6810 physicians met the inclusion criteria. Nineteen cognitive biases were identified.

**Results:** All studies found at least one cognitive bias or personality trait to affect physicians. Overconfidence, lower tolerance to risk, the anchoring effect, and information and availability biases were associated with diagnostic inaccuracies in 36.5 to 77 % of case-scenarios. Five out of seven (71.4 %) studies showed an association between cognitive biases and therapeutic or management errors. Of two (10 %) studies evaluating the impact of cognitive biases or personality traits on patient outcomes, only one showed that higher tolerance to ambiguity was associated with increased medical complications (9.7 % vs 6.5 %; $p = .004$). Most studies (60 %) targeted cognitive biases in diagnostic tasks, fewer focused on treatment or management (35 %) and on prognosis (10 %). Literature gaps include potentially relevant biases (e.g. aggregate bias, feedback sanction, hindsight bias) not investigated in the included studies. Moreover, only five (25 %) studies used clinical guidelines as the framework to determine diagnostic or treatment errors. Most studies ($n = 12$, 60 %) were classified as low quality.

**Conclusions:** Overconfidence, the anchoring effect, information and availability bias, and tolerance to risk may be associated with diagnostic inaccuracies or suboptimal management. More comprehensive studies are needed to determine the prevalence of cognitive biases and personality traits and their potential impact on physicians' decisions, medical errors, and patient outcomes.

**Keywords:** Decision making, Cognitive bias, Personality traits, Cognition, Physicians, Case-scenarios, Systematic review

* Correspondence: gustavo.saposnik@econ.uzh.ch
†Equal contributors
[1]Department of Economics, University of Zurich, Zürich, Switzerland
[2]Stroke Program, Department of Medicine, St Michael's Hospital, University of Toronto, Toronto M5C 1R6, Canada
Full list of author information is available at the end of the article

## Background

Medical errors occur in 1.7-6.5 % of all hospital admissions causing up to 100,000 unnecessary deaths each year, and perhaps one million in excess injuries in the USA [1, 2]. In 2008, medical errors cost the USA $19.5 billion [3]. The incremental cost associated with the average event was about US$ 4685 and an increased length of stay of about 4.6 days. The ultimate consequences of medical errors include avoidable hospitalizations, medication underuse and overuse, and wasted resources that may lead to patients' harm [4, 5].

Kahneman and Tversky introduced a dual-system theoretical framework to explain judgments, decisions under uncertainty, and cognitive biases. System 1 refers to an automatic, intuitive, unconscious, fast, and effortless or routine mechanism to make most common decisions (Fig. 1). Conversely, system 2 makes deliberate decisions, which are non-programmed, conscious, usually slow and effortful [6]. It has been suggested that most cognitive biases are likely due to the overuse of system 1 or when system 1 overrides system 2 [7–9]. In this framework, techniques that enhance system 2 could counteract these biases and thereby improve diagnostic accuracy and decrease management errors.

Concerns about cognitive biases are not unique to medicine. Previous studies showed the influence of cognitive biases on decisions inducing errors in other fields (e.g., aeronautic industry, factory production) [10, 11]. For example, a study investigating failures and accidents identified that over 90 % of air traffic control system errors, 82 % of production errors in an unnamed company, and 50–70 % of all electronic equipment failures were partly or wholly due to human cognitive factors [10]. Psychological assessments and quality assessment tools (e.g. Six Sigma) have been applied in many sectors to reduce errors and improve quality [12–15].

The health sector shares commonalities with industrial sectors including vulnerability to human errors [11, 14]. Therefore, a better understanding of the available evidence on cognitive biases influencing

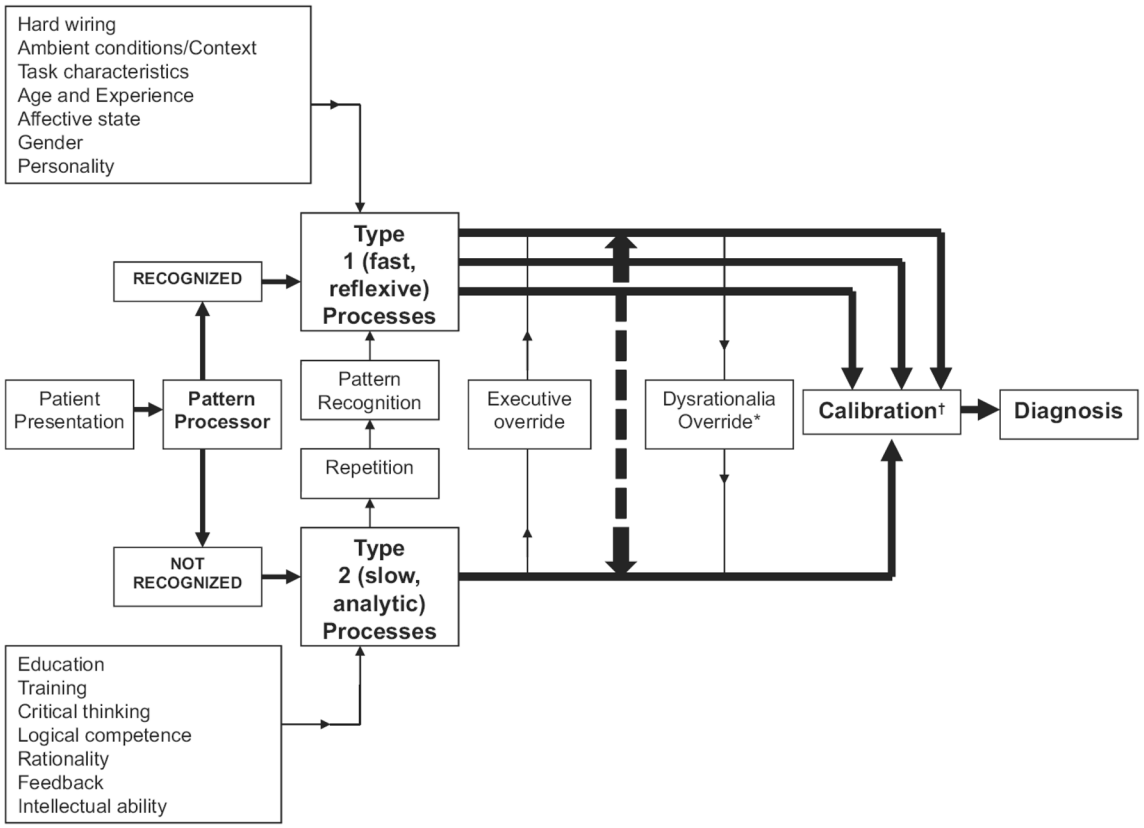



**Fig. 1** A model for diagnostic reasoning based on dual-process theory (from Ely et al. with permission).[9] System 1 thinking can be influenced by multiple factors, many of them subconscious (emotional polarization toward the patient, recent experience with the diagnosis being considered, specific cognitive or affective biases), and is therefore represented with multiple channels, whereas system 2 processes are, in a given instance, single-channeled and linear. System 2 overrides system 1 (executive override) when physicians take a time-out to reflect on their thinking, possibly with the help of checklists. In contrast, system 1 may irrationally override system 2 when physicians insist on going their own way (e.g., ignoring evidence-based clinical decision rules that can usually outperform them). Notes: Dysrationalia denotes the inability to think rationally despite adequate intelligence. "Calibration" denotes the degree to which the perceived and actual diagnostic accuracy correspond

medical decisions is crucial. Such an understanding is particularly needed for physicians, as their errors can be fatal and very costly. Moreover, such an understanding could also be useful to inform learning strategies to improve clinical performance and patient outcomes, whereas literature gaps could be useful to inform future research.

In the last three decades, we learned about the importance of patient- and hospital-level factors associated with medical errors. For example, standardized approaches (e.g. Advanced Trauma Life Support, ABCs for cardiopulmonary resuscitation) at the health system level lead to better outcomes by decreasing medical errors [16, 17]. However, physician-level factors were largely ignored as reflected by reports from scientific organizations [18–20]. It was not until the 1970s that cognitive biases were initially recognized to affect individual physicians' performance in daily medical decisions [6, 21–24]. Despite these efforts, little is known about the influence of cognitive biases and personality traits on physicians' decisions that lead to diagnostic inaccuracies, medical errors or impact on patient outcomes. While a recent review on cognitive biases and heuristics suggested that general medical personnel is prone to show cognitive biases, it did not answer the question whether these biases actually relate to the number of medical errors in physicians [25].

In the present (primarily narrative) systematic review, we therefore reviewed the literature reporting the existing evidence on the relation between cognitive biases affecting physicians and medical decisions. Under the concept of cognitive biases, we also included personality traits (e.g. aversion to risk or ambiguity) that may systematically affect physicians' judgments or decisions, independent of whether or not they result in immediate medical errors. Over 32 types of cognitive biases have been described [26]. Importantly, some of these may reflect personality traits that could result in choice tendencies that are factually wrong, whereas others reflect decisions that are potentially suboptimal, although there is no objectively "correct" decision (e.g. risk aversion, tolerance to ambiguity). Both of these factors were included here.

Our review has four objectives: 1) to identify the most common cognitive biases by subjecting physicians to real world situations or case-vignettes, 2) to evaluate the influence of cognitive biases on diagnostic accuracy and medical errors in management or treatment, 3) to determine which cognitive biases have the greatest impact on patient outcomes, and 4) to identify literature gaps in this specific area to guide future research. After addressing these objectives, we conclude by highlighting the practical implications of our findings and by outlining an action plan to advance the field.

## Methods

### Data sources

We conducted a literature search of MEDLINE and the Cochrane Library databases from 1980 to May 2015 by using a pre-specified search protocol (Additional file 1). We used a permuted combination of MeSH terms as major subjects, including: "medical errors", "bias", "cognition", "decision making", "physicians", and "case-vignettes" or "case-scenarios". In-line with the learning and education literature, case-vignettes, clinical scenarios or 'real world' encounters are regarded as the best simple strategy to evaluate cognitive biases among physicians [27]. In addition, this approach has also the advantage of facilitating the assessment of training strategies to ameliorate the influence of cognitive biases on medical errors. We therefore restricted our sample to studies that used case-vignettes or real-world encounters.

Results of the combination of search terms are listed in the Additional file 1. We also completed further searches based on key words, and reviewed references from previously retrieved articles. All articles were then combined into a single list, and duplicates ($n = 106$) were excluded (Fig. 2).

#### *Study selection*

Candidate articles examining cognitive biases influencing medical decisions were included for review if they met the following five inclusion criteria: First, the study was conducted on physicians. Second, at least one outcome measure was reported. Third, at least one cognitive factor or bias was investigated and defined a priori. Fourth, case-vignettes or real clinical encounters were used [28]. Fifth, the study was written in English. We analyzed the number of articles that fulfilled our inclusion criteria on each cognitive factor or bias, methodological aspects, and the magnitude of effect (as prevalence or odds ratios) on diagnostic or therapeutic decisions. We excluded studies that were not the primary source. We analyzed the original data as reported by the authors. Studies not providing raw data were also excluded (e.g. review articles, letters to Editors).

A recent systematic review was focused on medical personnel in general rather than physicians, and therefore included a different set of studies in their analysis than those that are of interest when considering the impact of cognitive biases on physicians' medical decision-making and medical errors (the focus of the current study) [25].

### Data extraction

We extracted data according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Fig. 2) [29]. Two reviewers (GS, librarian) assessed titles and abstracts to determine

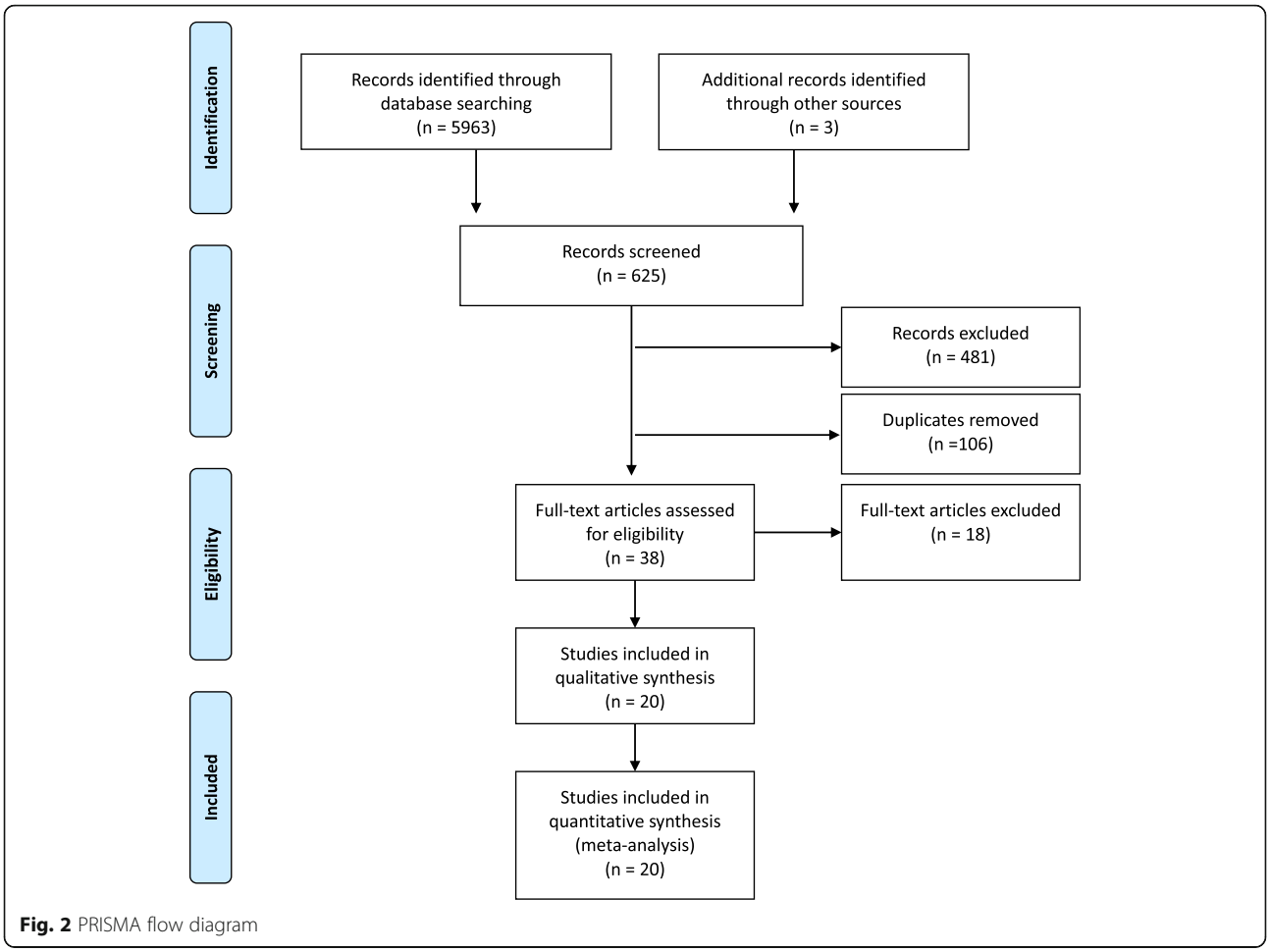



**Fig. 2** PRISMA flow diagram

eligibility. Data were extracted using standardized collection forms. Information was collected on country of origin, study design, year of publication, number of studied cognitive biases, population target (general practitioners, specialists, residents), decision type (e.g. diagnosis, treatment, management), unadjusted vs. adjusted analysis (for measured confounders, such as age, years of training, expertise), type of outcome (see below), data quality, and summary main findings. We also included descriptive elements (attributes) of the medical information provided for each case-scenario. The main outcomes were any form of medical error [26, 30], including: underuse or overuse of medical tests, diagnostic accuracy, lack of prescription or prescription of unnecessary medications, outcomes of surgical procedures, and avoidable hospitalizations.

### *Data quality*

We used the Newcastle-Ottawa Scale (NOS) to assess the quality of studies (see Additional file 2) [31]. The NOS is a quality assessment tool for observational studies recommended by the Cochrane Collaboration [32]. It assigns one or two points for each of eight items, categorized into three groups: the selection of the study groups; the comparability of the groups; and the ascertainment of the outcome of interest. Previous studies defined NOS scores as: 7–9 points considered as high quality, 5–6 as moderate quality, and 0–4 as low quality [33]. For example, studies that do not provide a description of the cohort, ascertainment of the exposure, adjustment for major confounders, or demonstration that the outcome of interest was not present at the beginning of the study were ranked as low quality [31].

## Results

We identified 5963 studies for the combination of MESH terms “decision making” and “physicians”. Of these, 114 fulfilled the selection criteria and were retrieved for detailed assessment. Among them, 38 articles used case-vignettes or real case scenarios in physicians (Fig. 2). Combinations of other search terms are shown in the Additional file 1: Table S1. Twenty studies comprising 6810 physicians (median 180 per study; range: 36–2206) met the inclusion criteria (Fig. 2) [30, 34–52].

In 55 % ($n = 11$) of the retained studies, results were adjusted for confounders, such as age, gender, level of training (see Additional file 1 for further details). Importantly, only five (25 %) studies used clinical guidelines as the framework to determine diagnostic or treatment errors, illustrating the scarcity of research on evidence-based decision making (e.g. GRADE: decisions based on levels of evidence provided by randomized trials, meta-analysis, etc).

#### Population target

Eight (40 %) studies included residents, six (30 %) studies included general practitioners, six (30 %) studies included internists, three (15 %) studies included emergency physicians and seven (35 %) studies included other specialists (Table 2). Ten (50 %) studies were conducted in the USA. Only six (30 %) studies classified errors based on real life measures, such as patient encounters, pathological images or endoscopic procedures, whereas the remaining 14 used narrative case-vignettes. Studies included a wide variety of medical situations, most commonly infections (upper respiratory tract, urinary tract) and cardiovascular disease (coronary disease, cerebrovascular disease) (Table 1). In summary, the included studies covered a wide range of medical conditions and participants.

#### Data quality

All studies were designed as cohort studies evaluating cognitive biases. According to the NOS, the majority of studies ($n = 12$, 60 %) were low quality, seven (35 %) studies ranked moderate and only one ranked as high quality [43] (see Additional file 2: Table S2 for details). All studies were classified as representative of the entire population (defined as how likely the exposed cohort was included in the population of physicians).

#### Presence of most common cognitive biases (Objective 1)

Our first objective was to evaluate the most common cognitive biases affecting physicians' decisions. Altogether, studies evaluated 19 different cognitive biases (Table 1 and Additional file 1).

It is important to bear in mind that these studies do not systematically assess each cognitive bias or personality traits. As a result, it is not possible to provide a true estimate of the prevalence of all cognitive biases among physicians. Overall, at least one cognitive factor or bias was present in all studies. Studies evaluating more than two cognitive biases, found that 50 to 100 % of physicians were affected by at least one [39, 50, 52]. Only three manuscripts evaluated more than 5 cognitive biases in the same study, in-line with the narrow scope of most studies [39, 50, 52]. One third of studies ($n = 6$) were descriptive, i.e., they provided the frequency of the cognitive bias without outcome data [36, 37, 39, 44, 48, 51].

The most commonly studied personality trait was tolerance to risk or ambiguity ($n = 5$), whereas the framing effects ($n = 5$) and overconfidence ($n = 5$) were the most common cognitive biases. There was a wide variability in the reported prevalence of cognitive biases (Fig. 3). For example, when analyzing the three most comprehensive studies that accounted for several cognitive biases (Fig. 4), the availability bias ranged from 7.8 to 75.6 % and anchoring from 5.9 to 87.8 %, suggestive of substantial heterogeneity among studies. In summary, cognitive biases may be common and present in all included studies. The framing effect, overconfidence, and tolerance to risk/ambiguity were the most commonly studied cognitive biases. However, methodological limitations make it difficult to provide an accurate estimation of the true prevalence.

#### Effect of cognitive biases on medical tasks (Objective 2)

Our second objective concerned the assessment of the influence of cognitive biases on diagnostic, medical management or therapeutic tasks. Most studies (12/20; 60 %) targeted cognitive biases in diagnostic tasks, 7 (35 %) studies targeted treatment or management tasks, and 2 studies (10 %) focused on errors in prognosis. The main measure was diagnostic accuracy in 35 % (7/20) of studies (Fig. 5). Overall, the presence of cognitive biases was associated with diagnostic inaccuracies in 36.5 to 77 % of case-scenarios [30, 35, 40, 42, 45, 52, 53]. A study including 71 residents, fellows, and attending pathologists evaluated 2230 skin biopsies with a diagnosis confirmed by a panel of expert pathologists. Information biases, anchoring effects, and the representativeness bias were associated with diagnostic errors in 51 % of 40 case-scenarios (compared to 16.4 % case-scenarios leading to incorrect diagnoses not related to cognitive biases; $p = 0.029$) [52].

Only seven (35 %) studies provided information to evaluate the association between physicians' cognitive biases and therapeutic or management errors [38, 41–43, 46, 47, 50]. Five out of the seven (71.4 %) studies showed an association between cognitive biases and these errors [38, 43, 46, 47, 50]. One study showed that overutilization of screening for prostate cancer among healthy individuals was associated with lower aversion to uncertainty ($p < 0.01$) [46]. In another study including 94 obstetricians who cared for 3488 deliveries, better coping strategies ($p < .015$) and tolerance to ambiguity ($p < .006$) were associated with optimal management (reflected by lower instrumental vaginal deliveries) and lower errors [43]. In a study including 32 anesthesiology residents, several cognitive biases (anchoring, overconfidence, premature closure, confirmation bias, etc.) were associated to errors in half of the 38 simulated encounters [50]. Two studies evaluating triage strategies for patients with bronchiolitis and coronary artery disease showed no

**Table 1** Characteristics of studies included in the systematic review

| Author | Year of publication | Country | Number participants | Methods | Clinical problem | Type of decision | Cognitive bias (n) | Type of cognitive bias | Data quality* |
|---|---|---|---|---|---|---|---|---|---|
| Redelmeier | 1995 | Canada | 639 | Survey | Ostoearthritis, TIA | Management and Treatment | 1 | Multiple alternative/Decoy bias | 5 |
| Ross | 1999 | UK | 407 | Survey | Depression | Treatment and management | 1 | Outcome bias | 6 |
| Graber | 2000 | USA | 232 | Survey | Headache, abdominal pain, depression | Diagnosis | 1 | Information bias | 4 |
| Sorum | 2003 | USA, France | 65 | Survey | Prostate cancer | Diagnosis | 1 | risk aversion | 4 |
| Baldwin | 2005 | USA | 46 | Experimental | Brochiolitis | Management | 2 | risk aversion, Ambiguity tolerance | 5 |
| Friedman | 2005 | USA | 216 | Survey | NR | Diagnosis | 1 | Overconfidence | 4 |
| Reyna | 2006 | USA | 74 | Survey | Unstable angina | Diagnosis and management | 1 | risk aversion | 5 |
| Bytzer | 2007 | Denmark | 127 | Video-cases | Reflux, epigastric pain | Diagnosis | 1 | Infromation bias | 4 |
| Dibonaventura | 2008 | USA | 2206 | Survey | Immunization | Treatment | 2 | omissions and naturalness bias | 4 |
| Mamede | 2010 | Netherlands | 36 | Experiment | Hepatitis, IBD, MI, Wernicke, Pneumonia, UTI, Meningitis | Diagnosis | 1 | Availability, Reflective reasoning | 5 |
| Mamade | 2010 | Netherlands | 84 | Survey | Aortic dissection, pancreatitis, hepatitis, pericarditis, hyperthiroidism, sarcoidosis, lung cancer, pneumonia, claudication, bacterial endocarditis | Diagnosis | 1 | Deliveration without attention | 3 |
| Gupta | 2011 | USA | 587 | Survey | Abdominal pain, headache, trauma, asthma, chest pain | Diagnosis | 1 | Outcome bias | 6 |
| Perneger | 2011 | Switzerland | 1439 | Survey | HIV infection | Treatment-Prognosis | 1 | Framing effect | 4 |
| Stiegler | 2012 | USA | 64 | Delphi and 38 simulated encounters | anaphylaxis, malignant hyperthermia, difficult airway, and pulmonary embolism | Treatment and management | 10 | anchoring, availability bias, premature closure, feedback bias, framing effect, confirmation bias, omission | 4 |
| Ogdie | 2012 | USA | 41 | Narratives | NR | Diagnosis | 9 | Anchoring, availability, framing effect, blind obedience, confirmation | 3 |
| Meyer | 2013 | USA | 118 | Survey | Abdominal pain, headache and rash, fever and arthralgias | Diagnosis | 1 | Overconfidence | 4 |
| Crowley | 2013 | International | 71 | Pathology cases | Vesicular and diffuse dermatitides | Diagnosis | 8 | anchoring, availability bias, confirmation bias, overconfidence | 4 |
| Saposnik | 2013 | Canada | 111 | Case-scenarios from real practice | Stroke | Prognosis | 2 | Overconfidence, anchoring | 5 |
| Msaouel | 2014 | Greece | 153 | Survey | Tuberculosis, CAD | Diagnosis | 2 | Gambler's and Conjunction fallacy | 5 |
| Yee | 2014 | USA | 94 | Experimental | Deliveries | Management and Treatment | 1 | Ambiguity tolerance/aversion | 7 |

*Data quality assessed using the Newcastle-Ottawa acale (NOS)

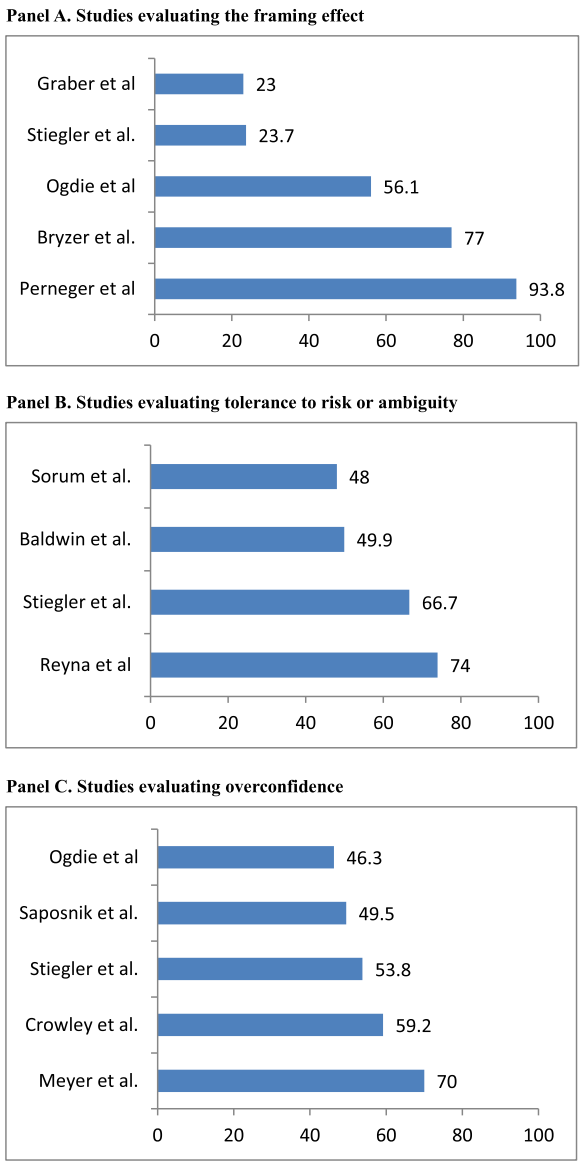



**Fig. 3** Prevalence of most common cognitive biases as reported by different studies. Numbers represent percentages reflecting the frequency of the cognitive factor/bias. Panel **a** represent the prevalence of the framing effect. Panel **b** represent the prevalence of prevalence of tolerance to risk and ambiguity. Panel **c** represents the prevalence of overconfidence. Overall, overconfidence and low tolerance to risk or ambiguity were found in 50-70 % of participants, whereas a wide variation was found for the framing effect

association between personality traits (e.g. risk aversion or tolerance to uncertainty) and hospital admissions [41, 42].

In summary, our findings suggest that cognitive biases (from one to two thirds of case-scenarios) may be associated with diagnostic inaccuracies. Evidence from five out of seven studies suggests a potential influence of cognitive biases on management or therapeutic errors [38, 43, 46, 47, 50]. Physicians who exhibited information bias, anchoring effects and representativeness bias, were more likely to make diagnostic errors [38, 43, 46, 50].

Further studies are needed to identify what the most common cognitive biases and the most effective strategies to overcome their potential influence of medical tasks and errors.

### Effect of physician's cognitive biases on patient outcomes (Objective 3)

The third objective of the present study was to determine the impact of cognitive biases on patient outcomes (e.g. avoidable hospitalizations, complications related to a procedure or medication, exposure to unnecessary invasive tests, etc). Only two (10 %) studies provided information to answer this question, both evaluating physicians' tolerance to uncertainty [41, 43]. In a study evaluating obstetrical practices, higher tolerance to ambiguity was associated with an increased risk of postpartum hemorrhage (9.7 % vs 6.5 %; $p = .004$). The negative effects persisted in the multivariable analysis (for postpartum hemorrhage: OR 1.51, 95 % CI 1.10–2.20 and for chorioamninitis: OR 1.37, 95 % CI 1.10–1.70) [43]. This phenomenon could be explained by overconfidence and underestimation of risk factors associated with maternal infections or puerperal bleeding. On the other hand, a study including 560 infants with bronchiolitis presented to the emergency department cared for by 46 pediatricians showed similar admission rates among physicians with low and high risk aversion or discomfort with diagnostic uncertainty (measured using a standardized tool) [41].

In summary, there too little evidence to make definitive conclusions on the influence of physicians' personality traits or cognitive biases on patient outcomes.

### Literature gaps and recommendations (Objective 4)

We systematically reviewed gaps in the literature. First, most of the studies (60 %) provided a qualitative definition of cognitive biases based on the interpretation of comments made by participants (e.g. illustrative quotes), lacking a unified and objective assessment tool [39, 50]. Second, the unit of study varies from study to study. For example, some authors report results based on the number of physicians involved in the study, whereas others report the results based on the number of case-scenarios. Third, limited information is currently available on the impact of cognitive biases on evidence-based care, as only 15 % of the studies were based on or supported by clinical guidelines (Table 2). Fourth, only one study evaluated the effect of an intervention (e.g. reflective reasoning) to ameliorate cognitive biases in physicians [35]. Fifth, most studies were classified as low quality according to NOS criteria. However, this scale is regarded as having a modest inter-rater reliability. We need consensus among researchers on the best tools to

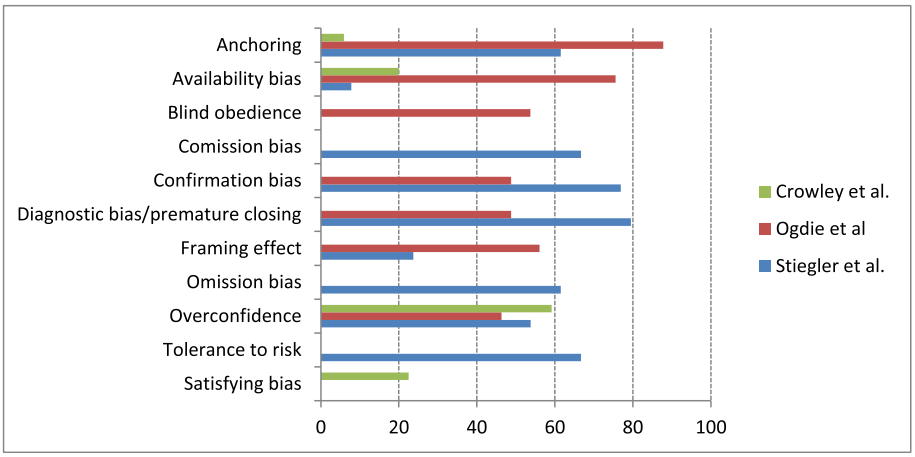



**Fig. 4** Prevalence of cognitive biases in the top three most comprehensive studies [39, 50, 52] Numbers represent percentages reflecting the frequency of the cognitive bias. Note the wide variation in the prevalence of cognitive biases across studies

assess the quality of manuscripts. Sixth, only two studies evaluated the influence of physicians' biases on patient outcomes. Finally, considering the great majority of studies (85 %) targeted only one or two biases (Table 1), the true prevalence of cognitive biases influencing medical decisions remains unknown.

As mentioned, medical errors are common in medical practice [5]. Physicians' biases and personality traits may explain, at least in part, some medical errors. Given the wide practice variability across medical disciplines, decisions on screening tests, surgical procedures, preventative medications, or other interventions (e.g. thrombolysis for acute stroke, antibiotics for an underlying infection, etc.) may not require the same cognitive abilities it is therefore likely that studies from one discipline cannot be transferred automatically to a different discipline. By extension, physicians' personality traits (e.g. aversion to ambiguity, tolerance to uncertainty) or cognitive biases (e.g. overconfidence) may not equally influence patient outcomes or medical errors in all disciplines. Time-urgency of the medical decision may be a relevant characteristic. Thus, a discipline-based research approach may be needed. There is scarce information in some disciplines and areas, including anesthesiology (decisions on procedures and anesthetic agents), emergency care, obstetrics and gynecology (e.g. decisions on procedures and primary care on women's health), endoscopic procedures (e.g. gastrointestinal, uropelvic), neurology (e.g. decision in multiple sclerosis and stroke care).

## Discussion

Early recognition of physicians' cognitive and biases are crucial to optimize medical decisions, prevent medical errors, provide more realistic patient expectations, and

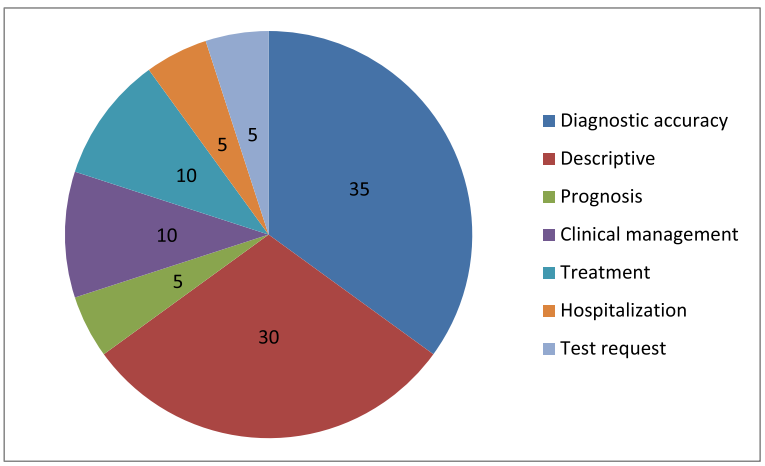



**Fig. 5** Outcome measures of studies evaluating cognitive biases. Numbers represent percentages. Total number of studies = 20. Note that 30 % of studies are descriptive and 35 % target diagnostic accuracy. Only few studies evaluated medical management, treatment, hospitalization or prognosis

**Table 2** Participants, attributes and outcomes of included studies

| Author | Type of participants | Number of vignettes or medical cases | Number of attributes | Based on Guidelines | Outcome measure | Type of outcome[a] | Type of analysis | Data quality[b] | Main findings |
|---|---|---|---|---|---|---|---|---|---|
| Redelmeier | GPs and Neurologist | 4 | 10-11 | yes | Treatment recommendations | 4 | unadjusted | 5 | Multiple options decreased the likelihood of medication prescription for pain and carotid endarterectomy by 26 % and 35 %, respectively |
| Ross | GPs | 3 | NA | No | Descriptive | 5 | adjusted | 6 | GPs were less likely to arrange a further consultation for female patients than for male patients (OR = 0.55). GPs with a pessimistic belief about depression were less likely to discuss non-physical symptoms or social factors; More experienced GPs were less likely to conduct a physical examination (OR = 0.60). |
| Graber | GPs | 2 | 8-9 | No | Descriptive | 1 | adjusted | 4 | GPs were less likely to believe a serious medical condition among patients with history of depression or somatic symptoms |
| Sorum | GPs | 32 | 5 | yes | Probability of ordering a test | 4 | adjusted | 4 | PSA were more likely ordered among GPs with discomfort for uncertainty and those who expressed regret. |
| Baldwin | Pediatric ED physicians | 397 | NA | No | Admission rates | 4 | adjusted | 5 | Risk aversion scores higher for physicians with >15 years of experience. Admissions rates did not differ between high and low risk adverse physicians (31.1 vs 30.1; p = 0.91). Adjusted admission rates did not different between high and low discomfort with uncertainty (32.3 vs 29.7; p = 0.84) |
| Friedmann | Medical students (72), residents (72), physicians (72) | 36 (9) | >20 | No | Diagnostic accuracy | 5 | adjusted | 4 | Overconfident found in 41 % of residents and in 36 % faculty. |
| Reyna | GPs and specialists | 9 | NA | Yes | Diagnostic accuracy and management | 6 | adjusted | 5 | Physicians deviated from Guidelines in terms of discharge. GP were more risk averse and less likely to discharge patients. Experts achieved better case-risk discrimination by processing less information |
| Bytzer | Specialists | 5 | NA | No | Diagnostic accuracy | 6 | unadjusted | 4 | Only 23 % endoscopists gave the same diagnosis for the two identical video-cases. The great majority were affected by prior information bias. |
| Dibonaventura | Physicians | 2 | 11–12 | No | Descriptive | 4 | unadjusted | 4 | Naturalness bias present in 40 %, omission bias in 60 % of participants |
| Mamede | Residents | 8 | NA | No, confirmed diagnosis | Diagnostic accuracy | 5 | unadjusted | 5 | Availability bias increased with years of training. Clinical reasoning ameliorate this bias |
| Mamade | internal medicine residents (34) and medical students (50) | 12 | >20 | No | Diagnostic accuracy | 6 | unadjusted | 3 | Conscious deliberation improved the likelihood of correct diagnosis in physicians, but not in medical students problems were complex, whereas reasoning mode did not matter in simple problems. In contrast, deliberation without attention improved novices' decisions. |

**Table 2** Participants, attributes and outcomes of included studies *(Continued)*

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Gupta | ED Physicians | 6 | >20 | No | Descriptive | 1 | adjusted | 6 | Outcome bias tends to inflate ratings in the presence of a positive outcome more than it penalizes scenarios with negative ones. |
| Perneger | GPs and specialists, and patients (1121) | 1 | 5 | No | Rating of new drug | 6 | adjusted | 4 | Physicians and patients provided higher value to the hypothetical new medication when presented in relative terms. Compared to descriptive information, relative mortality reduction (OR 4.40; 3.05 – 6.34), Number needed to treat (OR 1.79; 1.21 – 2.66), and relative survival extension (OR 4.55; 2.74 – 7.55) had a more positive perception. |
| Stiegler | Residents (32), Faculty (32) | 20 | NA | Catalogue of common cases | Management | 1 | unadjusted | 4 | 1. Developed a cognitive factor/bias catalogue, 2. Top 10 cognitive biases and personality traits: anchoring, availability bias, omission bias, commission bias, premature closure, confirmation bias, framing effect, overconfidence, feedback bias, and sunk cost.<br>3. Errors perceived by faculty to be important to anesthesiology were indeed observed frequently among trainees in a simulated environment. |
| Ogdie | Residents | 41 | NA | No | Descriptive | 6 | unadjusted | 3 | Most common biases: anchoring (88 %), availability (76 %), framing effect (56 %), overconfidence (46 %) |
| Meyer | Physicians | 4 | 6-11 | No | Diagnostic accuracy | 2 | unadjusted | 4 | Higher confidence was related to decreased requests for additional diagnostic tests (P = .01); higher case difficulty was related to more requests for additional reference materials (P = .01). |
| Crowley | pathology residents, fellows and staff pathologists | 40 | NA | No | Diagnostic accuracy | 6 | unadjusted | 4 | Overall, biases occurred in 52 % of incorrect cases compared to 21 % correct. Most common biases-Availability (20 %) and satisfying biases (22.5 %) the two most common. All the rest, less than 10 %. |
| Saposnik | Residents, internists, emergency physicians and Neurologist | 10 | 5-7 | No | Probability of death or disability | 6 | adjusted | 5 | Higher confidence was not associated with better outcome predictions. 70 % of underestimated the risk of the death or disability, 38 % overestimated death at 30 days. |
| Msaouel | Residents | 2 | 4, 5 | No | Descriptive | 1 | adjusted | 5 | Gambler's fallacy in 46 %, conjunction bias 69 % |
| Yee | Specialists (Obstetricians) | 3488 | NA | No | Management | 6 | adjusted | 7 | Physicians with a higher tolerance of ambiguity were less likely to deliver patients by operative vaginal delivery (11.8 % vs 16.4 %; p = 0.006). The effect disappeared in the adjusted analysis (OR 0.77, 95 % CI 0.53-1.1) |

*NA* not available, *GP* general practitioners

[a]Type of outcome measured: 1 = probability, 2 = rating, 3 = ranking, 4 = yes/no choice, 5 = discrete choice, 6 = more than 2 alternatives

[b]Data quality assessed by the Newcastle-Ottawa Score. See details in the text and Additional file 2

contribute to decreasing the rising health care costs altogether [3, 8, 54]. In the present systematic review, we had four objectives. First, we identified the most commonly reported cognitive biases (i.e., anchoring and framing effects, information biases) and personality traits (e.g. tolerance to uncertainty, aversion to ambiguity) that may potentially affect physicians' decisions. All included studies found at least one cognitive factor/bias, indicating that a large number of physicians may be possibly affected [39, 50, 52]. Second, we identified the effect of physician's cognitive biases or personality traits on medical tasks and on medical errors. Studies evaluating physicians' overconfidence, the anchoring effect, and information or availability bias may suggest an association with diagnostic inaccuracies [30, 35, 40, 42, 45, 52, 53]. Moreover, anchoring, information bias, overconfidence, premature closure, representativeness and confirmation bias may be associated with therapeutic or management errors [38, 43, 46, 47, 50]. Misinterpretation of recommendations and lower comfort with uncertainty were associated with overutilization of diagnostic tests [46]. Physicians with better coping strategies and tolerance to ambiguity could be related to optimal management [43].

For our third objective – identifying the relation between physicians' cognitive biases and patient's outcomes- we only had very sparse data: Only 10 % of studies provided data on this area [41, 43]. Only one study showed higher complications (OR 1.51, 95 % CI 1.10–2.20) among patients cared for by physicians with higher tolerance to ambiguity [43]. The fourth and final objective was to identify gaps in the literature. We found that only few (<50 %) of an established set of cognitive biases [26] were assessed, including: overconfidence, and framing effects. Other listed and relevant biases were not studied (e.g. aggregation bias, feedback sanction, hindsight bias). For example, aggregation bias (the assumption that aggregated data from clinical guidelines do not apply to their patients) or hindsight bias (the tendency to view events as more predictable than they really are) both compromise a realistic clinical appraisal, which may also lead to medical errors [18, 26]. More importantly, only 35 % of studies provided information on the association between cognitive biases or personality traits and medical errors [38, 41–43, 46, 47, 50], with scarce information on their impact on patient outcomes, preventing us from making definite conclusions [41, 43]. Furthermore, the quality of the included studies was classified as low to modest according to NOS criteria, as most studies provided limited descriptions of the exposure and research cohort, and none contributed with follow-up data (e.g. sustainability and reliability of the effects or long-term outcomes) (Additional file 2).

When comparing the previous systematic review on patients and medical personnel [25] with ours, some commonalities are apparent. Both reviews agree on the relevance of the topic, identify that a systematic analysis of the impact of cognitive biases on medical decisions is lacking despite substantial work completed in the last two decades [25]. Having a different objective, the authors nicely summarized the number of studies that investigated each cognitive bias either in patients or medical personnel [25]. Similarly, cognitive biases seem to be common among physicians as identified in 80 % ($n = 51$) of studies included in Blumenthal-Barby and Krieger's review and all selected studies ($n = 20$) evaluating at least one outcome in the present review [25].

However, both studies were not able to provide an accurate estimate of the true prevalence of cognitive biases or personality traits affecting medical decisions in physicians.

On the other hand, our study adds relevant information regarding the influence of cognitive biases particularly in physicians on diagnostic inaccuracies, suboptimal management and therapeutic errors, and patient outcomes. Our first objective allowed the identification of additional biases (e.g. framing effect, decoy effect, default bias) or physician's personality traits (e.g. low tolerance to uncertainty, aversion to ambiguity), by including 14 further studies. We also completed a systematic quality assessment of each study using a standardized tool and identified gaps related to the influence of cognitive biases on medical errors [31].

#### What can be done?

The identification and recognition of literature gaps constitute the first step to finding potential solutions. Increasing awareness among physicians and medical students is an important milestone. A comprehensive narrative review comprising 41 studies on cognitive interventions to reduce misdiagnosis found three main effective strategies: increasing knowledge and expertise, improving clinical reasoning, and getting help from colleagues, experts and tools [55]. First, reflective reasoning counteracts the impact of cognitive biases by improving diagnostic accuracy in second- (OR 2.03; 95 % CI, 1.49–2.57) and first-year residents [OR (odds ratio) 2.31; 95 % CI, 1.89–2.73] [35]. Second, the implementation of tools (e.g. cognitive checklist, calibration) may overcome overconfidence, the anchoring and framing effects (Fig. 5) [8, 9, 56]. Third, heuristics approaches (shortcuts to ignore less relevant information to overcome the complexity of some clinical situations) can improve decision making. As shown by Marewski and Gigerenzer, the identification of three rules (search for predictors to determine their individual importance, stop searching when relevant information was already obtained, and a criteria that specifies how a decision is made) may facilitate prompt decisions and may help physicians to avoid errors in some clinical situations [21, 57, 58].

The inclusion of training in cognitive biases in graduate and postgraduate programs might foster medical education and thereby improve health care delivery [59].

A commitment from academic institutions, scientific organizations, universities, the public, and policy-makers would be needed to reduce a defensive medical practice [60, 61]. An initial step towards this goal may be the 'Choosing wisely' strategy [62, 63].

### What are the practical implications of our findings?

As shown, cognitive biases and personality traits may affect our clinical reasoning processes which may lead to errors in the diagnosis, management, or treatment of medical conditions [6, 26]. Errors perceived by faculty to be relevant were indeed observed in 50–80 % of trainees in real practice [50]. Misdiagnosis, mismanagement, and mistreatment are frequently associated with poorer outcomes, which are the most common reasons for patients' dissatisfaction and medical complaints [54, 64, 65].

Our study has several limitations that deserve comment. First, although we aimed to be as systematic as possible in reviewing the literature, we cannot rule out involuntary omissions. It is also possible that our results may be somewhat limited by the strictness of our inclusion criteria. Second, we were not able to complete a formal meta-analysis due to the diversity of definitions and data reported, and small number of studies evaluating specific cognitive biases. In particular, a limited number of studies evaluated the same constructs. Moreover, across studies we often found a lack (in 30 % of studies) or heterogeneity in the outcome measures, mixed denominators (some studies report their findings based on the number of participants, while others based on case-scenarios) [41, 43, 52], and different scope (e.g. some studies are descriptive, [36, 37, 39, 44, 48, 51] whereas others [7, 30, 35, 42, 43, 47, 50, 52] target diagnostic or therapeutic errors). Third, most studies use hypothetical case-vignettes which may not truly reflect medical decisions in real life. Fourth, the assessment of the number of medical elements included in each case scenario may not be consistent (some were reported by authors and others estimated based on the description of case-scenarios) [35, 40, 51]. Fifth, the use of the NOS to assess the quality of studies has been criticized for having modest inter-rater reliability [66, 67].

Despite the aforementioned limitations, our study reflects the relevance and potential burden of the problem, how little we know about the implications of cognitive biases and personality traits on physicians' decisions, and their impact on patients-oriented outcomes. Our findings may also increase physicians' awareness of own personality traits or cognitive biases when counseling or advising patients and their family members that may lead to medical errors. From a health policy perspective, this information would provide additional insights on medically relevant cognitive biases and personality traits that contribute the rising health care costs [3, 68].

## Conclusions

In the present systematic review, we highlighted the relevance of recognizing physicians' personality traits and cognitive biases. Although cognitive biases may affect a wide range of physicians (and influence diagnostic accuracy, management, and therapeutic decisions), their true prevalence remains unknown.

Thus, substantial gaps limit our understanding of the impact of cognitive biases on medical decisions. As a result, new research approaches are needed. We propose the design of more comprehensive studies to evaluate the effect of physicians' personality traits and biases on medical errors and patient outcomes in real medical encounters and interventions or using guideline-based case-vignettes. This can be accomplished by identifying physician characteristics, combining validated surveys and experiments commonly used in behavioral economics to elicit several critical personality traits (e.g. tolerance to uncertainty, aversion to risk and ambiguity), and cognitive biases (e.g. overconfidence, illusion of control). Prospective studies evaluating and comparing different training strategies for physicians are needed to better understand and ameliorate the potential impact of cognitive biases on medical decisions or errors. In addition, effective educational strategies are also needed to overcome the effect of cognitive biases on medical decisions and interventions. Together, this information would provide new insights that may affect patient outcomes (e.g. avoidable hospitalizations, complications related to a procedure or medication, request of unnecessary tests, etc) and help attenuate medical errors [3, 68, 69].

## Additional files

**Additional file 1:** Literature search and definitions of cognitive biases and personality traits. (PDF 101 kb)

**Additional file 2:** Newcastle-Ottawa assessment tool and data quality. (PDF 76 kb)

#### Abbreviations

CI: Confidence intervals; NOS: Newcastle-Ottawa Scale; OR: Odds ratio; PRISMA: Preferred reporting items for systematic reviews and meta-analyses

#### Acknowledgments

The authors thank Maria Terzaghi and Lauren Cipriano for her input and comments regarding the manuscript and methods of this systematic review. We are also grateful to David Lighfoot (Librarian at the Li Ka Shing Institute, University of Toronto) for his help with the literature search and data quality assessment.

#### Funding

None.

#### Availability of data and materials

All data is available in the manuscript and additional files 1 and 2.

#### Authors' contribution

GS participated in the conception, design, literature search, analysis, interpretation of the results, drafting the manuscript and made critical revisions of the manuscript. DR participated in the conception, design, interpretation of the results, drafting the manuscript and made critical

revisions of the manuscript. CCR participated in the conception, design, interpretation of the results, drafting the manuscript and made critical revisions of the manuscript. PNT participated in the conception, design, analysis, interpretation of the results, drafting the manuscript and made critical revisions of the manuscript. All authors read and approved the final manuscript.

#### Authors' information

Dr. Saposnik is supported by the Distinguished Clinician-Scientist Award from the Heart and Stroke Foundation of Canada (HSFC) following a peer-review and open competition. Dr. Redelmeier holds a Canada Research Chair in Decision making. Drs. Ruff and Tobler are supported by the Swiss National Science Foundation (CRSII3_141965).

#### Competing interests

The authors declare that they have no competing interests.

#### Consent for publication

Not applicable.

#### Ethics approval and consent to participate

Not applicable.

#### Financial support

None.

#### Author details

[1]Department of Economics, University of Zurich, Zürich, Switzerland. [2]Stroke Program, Department of Medicine, St Michael's Hospital, University of Toronto, Toronto M5C 1R6, Canada. [3]Institute for Clinical Evaluative Sciences (ICES), Toronto, Canada. [4]University of Zurich, 9 Blumplistrasse, Zurich (8006), Switzerland.

Received: 5 June 2016 Accepted: 25 October 2016
Published online: 03 November 2016

#### References

1. Classen DC, Pestotnik SL, Evans RS, Burke JP. Computerized surveillance of adverse drug events in hospital patients. JAMA. 1991;266(20):2847–51.
2. Bates DW, Cullen DJ, Laird N, Petersen LA, Small SD, Servi D, Laffel G, Sweitzer BJ, Shea BF, Hallisey R, et al. Incidence of adverse drug events and potential adverse drug events. Implications for prevention. ADE Prevention Study Group. JAMA. 1995;274(1):29–34.
3. Andel C, Davidow SL, Hollander M, Moreno DA. The economics of health care quality and medical errors. J Health Care Finance. 2012;39(1):39–50.
4. OECD. Health at a Glance 2013: OECD Indicators, OECD Publishing. 2013. http://dx.doi.org/10.1787/health_glance-2013-en.
5. Ioannidis JP, Lau J. Evidence on interventions to reduce medical errors: an overview and recommendations for future research. J Gen Intern Med. 2001;16(5):325–34.
6. Tversky A, Kahneman D. Judgment under Uncertainty: Heuristics and Biases. Science. 1974;185(4157):1124–31.
7. Mamede S, van Gog T, van den Berge K, van Saase JL, Schmidt HG. Why do doctors make mistakes? A study of the role of salient distracting clinical features. Acad Med. 2014;89(1):114–20.
8. van den Berge K, Mamede S. Cognitive diagnostic error in internal medicine. Eur J Intern Med. 2013;24(6):525–9.
9. Ely JW, Graber ML, Croskerry P. Checklists to reduce diagnostic errors. Acad Med. 2011;86(3):307–13.
10. Dhillon BS. Human errors: a review. Microelectron Reliab. 1989;29(3):299–304.
11. Stripe SC, Best LG, Cole-Harding S, Fifield B, Talebdoost F. Aviation model cognitive risk factors applied to medical malpractice cases. J Am Board Fam Med. 2006;19(6):627–32.
12. Chassin MR. Is health care ready for Six Sigma quality? Milbank Q. 1998;76(4):565–91. 510.
13. Corn JB. Six sigma in health care. Radiol Technol. 2009;81(1):92–5.
14. Zeltser MV, Nash DB. Approaching the evidence basis for aviation-derived teamwork training in medicine. Am J Med Qual. 2010;25(1):13–23.
15. Ballard S-B. The U.S. commercial air tour industry: a review of aviation safety concerns. Aviat Space Environ Med. 2014;85(2):160–6.
16. Kern KB, Hilwig RW, Berg RA, Sanders AB, Ewy GA. Importance of continuous chest compressions during cardiopulmonary resuscitation: improved outcome during a simulated single lay-rescuer scenario. Circulation. 2002;105(5):645–9.
17. Collicott PE, Hughes I. Training in advanced trauma life support. JAMA. 1980;243(11):1156–9.
18. Michaels AD, Spinler SA, Leeper B, Ohman EM, Alexander KP, Newby LK, Ay H, Gibler WB, American Heart Association Acute Cardiac Care Committee of the Council on Clinical Cardiology, Outcomes Research, et al. Medication errors in acute cardiovascular and stroke patients: a scientific statement from the American Heart Association. Circulation. 2010;121(14):1664–82.
19. Khoo EM, Lee WK, Sararaks S, Abdul Samad A, Liew SM, Cheong AT, Ibrahim MY, Su SH, Mohd Hanafiah AN, Maskon K, et al. Medical errors in primary care clinics–a cross sectional study. BMC Fam Pract. 2012;13:127.
20. Jenkins RH, Vaida AJ. Simple strategies to avoid medication errors. Fam Pract Manag. 2007;14(2):41–7.
21. Marewski JN, Gigerenzer G. Heuristic decision making in medicine. Dialogues Clin Neurosci. 2012;14(1):77–89.
22. Wegwarth O, Schwartz LM, Woloshin S, Gaissmaier W, Gigerenzer G. Do physicians understand cancer screening statistics? A national survey of primary care physicians in the United States. Ann Intern Med. 2012;156(5):340–9.
23. Elstein AS. Analytic methods and medical education. Problems and prospects. Med Decis Making. 1983;3(3):279–84.
24. Elstein AS. Clinical judgment: psychological research and medical practice. Science. 1976;194(4266):696–700.
25. Blumenthal-Barby J, Krieger H. Cognitive biases and heuristics in medical decision making: a critical review using a systematic search strategy. Med Decis Making. 2015;35(4):539–57.
26. Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. Acad Med. 2003;78(8):775–80.
27. Peabody JW, Luck J, Glassman P, Jain S, Hansen J, Spell M, Lee M. Measuring the quality of physician practice by using clinical vignettes: a prospective validation study. Ann Intern Med. 2004;141(10):771–80.
28. Eva KW. What every teacher needs to know about clinical reasoning. Med Educ. 2005;39(1):98–106.
29. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Intern Med. 2009;151(4):W65–94.
30. Meyer AN, Payne VL, Meeks DW, Rao R, Singh H. Physicians' diagnostic accuracy, confidence, and resource requests: a vignette study. JAMA Intern Med. 2013;173(21):1952–8.
31. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. [In http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp]
32. Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011] [In http://www.cochrane-handbook.org]
33. Shi Q, MacDermid J, Santaguida L, Kyu HH: Predictors of surgical outcomes following anterior transposition of ulnar nerve for cubital tunnel syndrome: A systematic review. J Hand Surg Am. 2011;36(12):1996–2001.e1–6. doi:10.1016/j.jhsa.2011.09.024.
34. Mamede S, Splinter TA, van Gog T, Rikers RM, Schmidt HG. Exploring the role of salient distracting clinical features in the emergence of diagnostic errors and the mechanisms through which reflection counteracts mistakes. BMJ Qual Saf. 2012;21(4):295–300.
35. Mamede S, van Gog T, van den Berge K, Rikers RM, van Saase JL, van Guldener C, Schmidt HG. Effect of availability bias and reflective reasoning on diagnostic accuracy among internal medicine residents. JAMA. 2010;304(11):1198–203.
36. Msaouel P, Kappos T, Tasoulis A, Apostolopoulos AP, Lekkas I, Tripodaki ES, Keramaris NC. Assessment of cognitive biases and biostatistics knowledge of medical residents: a multicenter, cross-sectional questionnaire study. Med Educ Online. 2014;19:23646.
37. Ross S, Moffat K, McConnachie A, Gordon J, Wilson P. Sex and attitude: a randomized vignette study of the management of depression by general practitioners. Br J Gen Pract. 1999;49(438):17–21.
38. Perneger TV, Agoritsas T. Doctors and patients' susceptibility to framing bias: a randomized trial. J Gen Intern Med. 2011;26(12):1411–7.
39. Ogdie AR, Reilly JB, Pang WG, Keddem S, Barg FK, Von Feldt JM, Myers JS. Seen through their eyes: residents' reflections on the cognitive and contextual components of diagnostic errors in medicine. Acad Med. 2012;87(10):1361–7.

40. Friedman C, Gatti G, Elstein A, Franz T, Murphy G, Wolf F. Are clinicians correct when they believe they are correct? Implications for medical decision support. Stud Health Technol Inform. 2001;84(Pt 1):454–8.
41. Baldwin RL, Green JW, Shaw JL, Simpson DD, Bird TM, Cleves MA, Robbins JM. Physician risk attitudes and hospitalization of infants with bronchiolitis. Acad Emerg Med. 2005;12(2):142–6.
42. Reyna VF, Lloyd FJ. Physician decision making and cardiac risk: effects of knowledge, risk perception, risk tolerance, and fuzzy processing. J Exp Psychol Appl. 2006;12(3):179–95.
43. Yee LM, Liu LY, Grobman WA. The relationship between obstetricians' cognitive and affective traits and their patients' delivery outcomes. Am J Obstet Gynecol. 2014;211(6):692 e691–696.
44. Graber MA, Bergus G, Dawson JD, Wood GB, Levy BT, Levin I. Effect of a patient's psychiatric history on physicians' estimation of probability of disease. J Gen Intern Med. 2000;15(3):204–6.
45. Bytzer P. Information bias in endoscopic assessment. Am J Gastroenterol. 2007;102(8):1585–7.
46. Sorum PC, Shim J, Chasseigne G, Bonnin-Scaon S, Cogneau J, Mullet E. Why do primary care physicians in the United States and France order prostate-specific antigen tests for asymptomatic patients? Med Decis Making. 2003;23(4):301–13.
47. Redelmeier DA, Shafir E. Medical decision making in situations that offer multiple alternatives. JAMA. 1995;273(4):302–5.
48. Dibonaventura M, Chapman GB. Do decision biases predict bad decisions? Omission bias, naturalness bias, and influenza vaccination. Med Decis Making. 2008;28(4):532–9.
49. Saposnik G, Cote R, Mamdani M, Raptis S, Thorpe KE, Fang J, Redelmeier DA, Goldstein LB. JURaSSiC: Accuracy of clinician vs risk score prediction of ischemic stroke outcomes. Neurol. 2013;81(5):448–55.
50. Stiegler MP, Ruskin KJ. Decision-making and safety in anesthesiology. Curr Opin Anaesthesiol. 2012;25(6):724–9.
51. Gupta M, Schriger DL, Tabas JA. The presence of outcome bias in emergency physician retrospective judgments of the quality of care. Ann Emerg Med. 2011;57(4):323–8. e329.
52. Crowley RS, Legowski E, Medvedeva O, Reitmeyer K, Tseytlin E, Castine M, Jukic D, Mello-Thoms C. Automated detection of heuristics and biases among pathologists in a computer-based system. Adv Health Sci Educ Theory Pract. 2013;18(3):343–63.
53. Mamede S, Schmidt HG, Rikers RM, Custers EJ, Splinter TA, van Saase JL. Conscious thought beats deliberation without attention in diagnostic decision-making: at least when you are an expert. Psychol Res. 2010;74(6):586–92.
54. Zwaan L, Thijs A, Wagner C, van der Wal G, Timmermans DR. Relating faults in diagnostic reasoning with diagnostic errors and patient harm. Acad Med. 2012;87(2):149–56.
55. Graber ML, Kissam S, Payne VL, Meyer AN, Sorensen A, Lenfestey N, Tant E, Henriksen K, Labresh K, Singh H. Cognitive interventions to reduce diagnostic error: a narrative review. BMJ Qual Saf. 2012;21(7):535–57.
56. Balla JI, Heneghan C, Glasziou P, Thompson M, Balla ME. A model for reflection for good clinical practice. J Eval Clin Pract. 2009;15(6):964–9.
57. Raab M, Gigerenzer G. The power of simplicity: a fast-and-frugal heuristics approach to performance science. Front Psychol. 2015;6:1672.
58. Elwyn G, Thompson R, John R, Grande SW. Developing IntegRATE: a fast and frugal patient-reported measure of integration in health care delivery. Int J Integr Care. 2015;15:e008.
59. Hyman DJ, Pavlik VN, Greisinger AJ, Chan W, Bayona J, Mansyur C, Simms V, Pool J. Effect of a physician uncertainty reduction intervention on blood pressure in uncontrolled hypertensives–a cluster randomized trial. J Gen Intern Med. 2012;27(4):413–9.
60. Kachalia A, Mello MM. Defensive medicine—legally necessary but ethically wrong?: Inpatient stress testing for chest pain in low-risk patients. JAMA Intern Med. 2013;173(12):1056–7.
61. Smith TR, Habib A, Rosenow JM, Nahed BV, Babu MA, Cybulski G, Fessler R, Batjer HH, Heary RF. Defensive medicine in neurosurgery: does state-level liability risk matter? Neurosurg. 2015;76(2):105–13. discussion 113-104.
62. Bhatia RS, Levinson W, Shortt S, Pendrith C, Fric-Shamji E, Kallewaard M, Peul W, Veillard J, Elshaug A, Forde I, et al. Measuring the effect of Choosing Wisely: an integrated framework to assess campaign impact on low-value care. BMJ Qual Saf. 2015;24(8):523–31. doi:10.1136/bmjqs-2015-004070.
63. Levinson W, Huynh T. Engaging physicians and patients in conversations about unnecessary tests and procedures: Choosing Wisely Canada. CMAJ. 2014;186(5):325–6.
64. Kallberg AS, Goransson KE, Ostergren J, Florin J, Ehrenberg A. Medical errors and complaints in emergency department care in Sweden as reported by care providers, healthcare staff, and patients - a national review. Eur J Emerg Med. 2013;20(1):33–8.
65. Studdert DM, Mello MM, Gawande AA, Gandhi TK, Kachalia A, Yoon C, Puopolo AL, Brennan TA. Claims, errors, and compensation payments in medical malpractice litigation. N Engl J Med. 2006;354(19):2024–33.
66. Hartling L, Milne A, Hamm MP, Vandermeer B, Ansari M, Tsertsvadze A, Dryden DM. Testing the Newcastle Ottawa Scale showed low reliability between individual reviewers. J Clin Epidemiol. 2013;66(9):982–93.
67. Lo CK, Mertz D, Loeb M. Newcastle-Ottawa Scale: comparing reviewers' to authors' assessments. BMC Med Res Methodol. 2014;14:45.
68. Stangierski A, Warmuz-Stangierska I, Ruchala M, Zdanowska J, Glowacka MD, Sowinski J, Ruchala P. Medical errors - not only patients' problem. Arch Med Sci. 2012;8(3):569–74.
69. Graber ML. The incidence of diagnostic error in medicine. BMJ Qual Saf. 2013;22 Suppl 2:ii21–7.

# B. Appendix to Study 2: i) Protocol

Saposnik *et al. BMC Neurology* (2016) 16:58
DOI 10.1186/s12883-016-0577-4

BMC Neurology

STUDY PROTOCOL Open Access

## Decision making under uncertainty, therapeutic inertia, and physicians' risk preferences in the management of multiple sclerosis (DIScUTIR MS)

Gustavo Saposnik[1,2*], Angel Perez Sempere[3], Roula Raptis[4], Daniel Prefasi[5], Daniel Selchen[1] and Jorge Maurino[5]

**Abstract**

**Background:** The management of multiple sclerosis (MS) is rapidly changing by the introduction of new and more effective disease-modifying agents. The importance of risk stratification was confirmed by results on disease progression predicted by different risk score systems. Despite these advances, we know very little about medical decisions under uncertainty in the management of MS. The goal of this study is to i) identify whether overconfidence, tolerance to risk/uncertainty, herding influence medical decisions, and ii) to evaluate the frequency of therapeutic inertia (defined as lack of treatment initiation or intensification in patients not at goals of care) and its predisposing factors in the management of MS.

**Methods/Design:** This is a prospective study comprising a combination of case-vignettes and surveys and experiments from Neuroeconomics/behavioral economics to identify cognitive distortions associated with medical decisions and therapeutic inertia. Participants include MS fellows and MS experts from across Spain. Each participant will receive an individual link using Qualtrics platform© that includes 20 case-vignettes, 3 surveys, and 4 behavioral experiments. The total time for completing the study is approximately 30–35 min. Case vignettes were selected to be representative of common clinical encounters in MS practice. Surveys and experiments include standardized test to measure overconfidence, aversion to risk and ambiguity, herding (following colleague's suggestions even when not supported by the evidence), physicians' reactions to uncertainty, and questions from the Socio-Economic Panel Study (SOEP) related to risk preferences in different domains. By applying three different MS score criteria (modified Rio, EMA, Prosperini's scheme) we take into account physicians' differences in escalating therapy when evaluating medical decisions across case-vignettes.

**Conclusions:** The present study applies an innovative approach by combining tools to assess medical decisions with experiments from Neuroeconomics that applies to common scenarios in MS care. Our results will help advance the field by providing a better understanding on the influence of cognitive factors (e.g., overconfidence, aversion to risk and uncertainty, herding) on medical decisions and therapeutic inertia in the management of MS which could lead to better outcomes.

This paper is published as Saposnik G, Sempere AP, Raptis R, Prefasi D, Selchen D, Maurino J. Decision making under uncertainty, therapeutic inertia, and physicians' risk preferences in the management of multiple sclerosis (DIScUTIR MS). BMC Neurol.,16, 58, 2016

Saposnik et al. BMC Neurology (2016) 16:58
DOI 10.1186/s12883-016-0577-4

BMC Neurology

STUDY PROTOCOL Open Access

# Decision making under uncertainty, therapeutic inertia, and physicians' risk preferences in the management of multiple sclerosis (DIScUTIR MS)

Gustavo Saposnik[1,2*], Angel Perez Sempere[3], Roula Raptis[4], Daniel Prefasi[5], Daniel Selchen[1] and Jorge Maurino[5]

## Abstract

**Background:** The management of multiple sclerosis (MS) is rapidly changing by the introduction of new and more effective disease-modifying agents. The importance of risk stratification was confirmed by results on disease progression predicted by different risk score systems. Despite these advances, we know very little about medical decisions under uncertainty in the management of MS. The goal of this study is to i) identify whether overconfidence, tolerance to risk/uncertainty, herding influence medical decisions, and ii) to evaluate the frequency of therapeutic inertia (defined as lack of treatment initiation or intensification in patients not at goals of care) and its predisposing factors in the management of MS.

**Methods/Design:** This is a prospective study comprising a combination of case-vignettes and surveys and experiments from Neuroeconomics/behavioral economics to identify cognitive distortions associated with medical decisions and therapeutic inertia. Participants include MS fellows and MS experts from across Spain. Each participant will receive an individual link using Qualtrics platform© that includes 20 case-vignettes, 3 surveys, and 4 behavioral experiments. The total time for completing the study is approximately 30–35 min. Case vignettes were selected to be representative of common clinical encounters in MS practice. Surveys and experiments include standardized test to measure overconfidence, aversion to risk and ambiguity, herding (following colleague's suggestions even when not supported by the evidence), physicians' reactions to uncertainty, and questions from the Socio-Economic Panel Study (SOEP) related to risk preferences in different domains. By applying three different MS score criteria (modified Rio, EMA, Prosperini's scheme) we take into account physicians' differences in escalating therapy when evaluating medical decisions across case-vignettes.

**Conclusions:** The present study applies an innovative approach by combining tools to assess medical decisions with experiments from Neuroeconomics that applies to common scenarios in MS care. Our results will help advance the field by providing a better understanding on the influence of cognitive factors (e.g., overconfidence, aversion to risk and uncertainty, herding) on medical decisions and therapeutic inertia in the management of MS which could lead to better outcomes.

* Correspondence: Gustavo.saposnik@econ.uzh.ch; saposnikg@smh.ca
[1]Division of Neurology, Department of Medicine, St. Michael's Hospital, University of Toronto, 55 Queen St E, Toronto, ON M5C 1R6, Canada
[2]Neuroeconomics and Decision Neuroscience, Department of Economics, University of Zurich, Zurich, Switzerland
Full list of author information is available at the end of the article

## Background

The field of multiple sclerosis (MS) has seen significant changes over the last several years [1, 2]. Clinicians and patients welcomed the introduction of disease-modifying therapy (DMT) for MS in the mid-1990s. Injectable agents, all with rather similar risk–benefit profiles, dominated MS care for over a decade. The approval of Natalizumab marked a change with the introduction of a more effective treatment option, but also entailed new risks associated with modulation of the immune system (e.g., risk of progressive multifocal leukoencephalopathy - PML) [2, 3]. More recently, the introduction of oral agents and new humanised monoclonal antibodies administered by infusion have opened yet another avenue for patients and clinicians [4]. Currently, there are over a dozen of DMTs available to treat MS, with varying availability around the world. Significant heterogeneity exists in the efficacy and risks associated with these therapies [5–7]. Therefore, clinicians have the challenge of tailoring treatment based on i) disease activity level (clinical and radiological data), ii) individual patient characteristics/preferences, iii) personal expertise/preference, in order to identify the optimal balance between efficacy and safety Table 1 (See Additional file 1 for data on some currently available agents) [8].

**Table 1** Comparative adverse events of different DMTs [7, 8]

| Disease modifying agent | Adverse events |
|---|---|
| Interferon beta | • Depression |
| | • thrombotic microangiopathy |
| | • hepatotoxicity |
| | • ISRs |
| | • Flu-like |
| | • LFT elevation |
| | • Leukopenia |
| Glatiramer acetate | • ISRs |
| | • Benign systemic reaction |
| Mitoxantrone | • Cardiac toxicity |
| | • Leukemia |
| Natalizumab | • Infusion reactions |
| | • PML |
| | • Infusion-related fatigue |
| Fingolimod | • Bradyarrhythmia |
| | • Macular edema |
| | • Herpes virus infection |
| | • PML |
| | • BCC |
| | • LFT elevation |
| | • Lymphopenia |
| | • Mild hypertension |
| Teriflunomide | • Hepatotoxicity |
| | • Peripheral neuropathy |
| | • Alopecia |
| | • Nausea/Diarrhea |
| Dimethyl fumarate | • Flushing |
| | • Gastrointestinal |
| | • PML |
| Alemtuzumab | • Infusion reactions |
| | • ITP |
| | • Goodpasture syndrome |
| | • Thyroid cancer |
| | • Infections |
| | • Autoimmune thyroid disease |

*ISRs* injection-site reactions, *LFT* liver function test, *PML* progressive multifocal leukoencephalopathy, *ITP* idiopathic thrombocytopenic purpura, *BCC* basal cell carcinoma

### Risk stratification in MS

An understanding of the risk of untreated multiple sclerosis is crucial to make therapeutic decisions Table 2 [8]. In addition, physicians' preferences and beliefs in effectiveness of treatment and drug safety profiles may influence their decisions. Disease activity/progression can be divided into physical, cognitive and radiological markers. Examples include number of attacks per year, number of disabling attacks, disability scales (clinical), lesion volume, GAD enhancing lesions, brain atrophy (MRI), and cognitive decline (e.g., using SDMT, PASAT, OR MoCA scales) [9]. Two scoring systems (Rio score and Modified Rio score) demonstrate good predictive value for MS progression. The Rio score includes MRI, clinical relapse and EDSS criteria, whereas the modified Rio score includes MRI and clinical relapse criteria (Fig. 1) [10]. A high risk profile using the modified Rio (score ≥2) includes either an MRI with more than 5 new T2 lesions (1 point) or 1 relapse in the first year (1 point) or two relapses within the first year of treatment (2 points) or the combination of these criterions [11]. These scores have been used to identify and predict response to Interferon β. For example, the modified Rio score in the PRISM trial revealed that participants who did not responded to Interferon β had a similar probability of disability progression as those assigned to the placebo group. Conversely, responders to Interferon β had a 52 % reduction in disability progression compared to placebo and non-responders ($p < 0.0001$). MS patients with a modified Rio score greater than or equal to 2 had a 65 % increased risk of disability progression (HR = 4.60; $p < 0.001$) [12]. A Canadian group concluded that a change in treatment may be considered in patients with relapsing remitting MS if there is a high level of concern in any one domain (relapses, progression or MRI), a medium level of concern in any two domains, or a low level of concern in all three domains [13]. The European

**Table 2** Risks of untreated relapsing MS

| Treatment targets | Evidence of association | Long-term outcome |
|---|---|---|
| T2 lesion volume | Increase of 0.8–I ml/year | Correlates with increased relapse frequency and long term disability outcomes. |
| T1 black hole conversion | 40–50 % of lesions go on to form black holes | Correlation with clinical measures and disability progression. |
| Brain atrophy | 0.5–1 %/year in MS vs. <0.1 % in healthy controls | Correlation with cognitive outcomes and EDSS in the long term. |
| Clinical relapses | Annualized relapse rate in placebo arms: 0.5–1.38 | Relapses associated with decreased quality of life. |
| | | Relapses associated with accrual of disability. |
| | | Earlier onset of SPMS. |
| Disability accrual | Average change of 0.27 EDSS points/per relapse | Increased likelihood of long term disability. |
| | MRI and lesional activity associated with disability progression | |

Reproduced with permission from Ontaneda et al. [8]

Medicines Agency approves escalating therapy with Natalizumab or Fingolimod in patients who had at least one relapse in the previous year while on Interferon β and either ≥9 T2-hyperintense lesions on brain MRI or ≥1 contrast-enhancing lesion MRI activity alone after the first year of treatment was associated with three- to fivefold increased risk of relapses or disability compared with stable patients. These recommendations have been supported by several prospective studies [14, 15].

Selection of a first line therapy will likely depend on several factors. Traditionally, and due to the availability of extended safety data, injectable agents may be the first choices. Given the comparable efficacy data between the injectable agents the selection of a therapy will be determined mostly by side effect profiles. Subjects with headaches, depression, and a history of liver dysfunction may experience worsening of these comorbidities when exposed to interferons. Monitoring for interferons includes following liver function tests, complete blood counts, and monitoring depression [8]. Given the availability of more effective drugs, the treatment paradigm is likely to change. However, it is expected there will be wide variability on the timing of this paradigm shift (e.g., starting more effective therapies as first line treatment) based on patients' and physicians' tolerance to risk, estimation of the clinical course, regional funding programs, among other factors. As a result, it is vital to identify situations for which physicians take the opportunity of escalating treatment when indicated (e.g., progression of disease determined by clinical relapses, EDSS disability score and imaging data).

### Therapeutic inertia: a new paradigm in MS

Therapeutic inertia is a term introduced in 2006 to define the lack of treatment initiation or intensification in patients not at goals of care [16–19]. Some examples include failure to intensify treatment in patients with persistent elevated blood pressure or blood glucose [16, 20, 21]. Reasons to explain therapeutic inertia include the lack of training and cultural organization in the practice at "treating to target", competing demands and clinical uncertainty [22, 23]. In the context of MS, therapeutic inertia is defined lack of treatment initiation or intensification when there is evidence of disease progression (based on clinical and radiological data). In the present study,

| Rio Score | | Modified Rio Score | |
|---|---|---|---|
| **Criterion** | **Change over the first year** | **Criterion** | **Change over the first year** |
| MRI criterion = 0<br>MRI criterion = 1 | ≤2 active* T2 lesions<br>>2 active T2 lesions | MRI criterion = 0<br>MRI criterion = 1 | ≤4 (5)‡ new T2 lesions<br>>4 (5)‡ new T2 lesions |
| Relapse criterion = 0<br>Relapse criterion = 1 | No relapses<br>≥1 relapse | Relapse criterion = 0<br>Relapse criterion = 1<br>Relapse criterion = 2 | No relapses<br>1 relapse<br>≥2 relapses |
| EDSS criterion = 0<br>EDSS criterion = 1 | Increase in EDSS score of <1 point<br>Increase in EDSS score of ≥1 point, sustained over at least 6 months | Not included | Not included |
| Rio Score = MRI criterion + relapse criterion + EDSS criterion | | Modified Rio Score = MRI criterion + relapse criterion | |

*Active lesions defined as new or enlarging T2-weighted lesions plus gadolinium-enhancing lesions over the first year. ‡The cut-off of four lesions was in the validation set; the cut-off of five lesions was in the training set. Abbreviation: EDSS, Expanded Disability Status Scale.

**Fig. 1** Modified with permission from Sormani et al. defining and scoring response to IFN-β in multiple sclerosis. Nat. Rev. Neurol. doi:10.1038/nrneurol.2013.146

disease progression was defined according to the modified Rio score, where patients had one or more recurrent attacks and/or an MRI with 5 or more new T2 lesions while receiving treatment with a disease- modifying agent [11]. Another more recent criterion strongly associated with risk of relapse or disability progression was the presence of isolated gadolinium-enhancing lesions [14, 15].

## Medical decision making

Making decisions in medical care is a complex task involving a variety of cognitive processes [24]. Decision making is defined as the process of examining possibilities, risks, uncertainties, and options, comparing them, and choosing a course of action [25, 26]. Decisions based on erroneous assessments may result in incorrect patient and family expectations, and potentially suboptimal advice, treatment, and prognosis. Moreover, many decisions are made with limited information from observational studies or clinical trials that may not apply to particular patients. Uncertainty is one of the most important reasons contributing to the status quo and making proactive therapeutic decisions [17, 23, 27]. We need a better understanding on how physicians decide about different therapeutic options under uncertainty for patients with MS.

## The problem

Despite the availability of different markers for risk stratification in patients with MS, it is difficult for expert clinicians to select the best strategy when the progression pattern of the disease is uncertain. MS experts and clinicians are trained to quickly recognize patterns or critical aspects of particular situations [28]. Some clinicians apply the knowledge they have acquired from previous experience, others use information available at the time of the assessment, others use risk score tools or a combination of the above. However, it is not known how MS experts behave in clinical scenarios with ambiguous outcomes (unknown probability or uncertain risk of an outcome) or when more therapeutic options become available. In addition, we have a limited understanding about physicians' beliefs and preferences on the widely available therapeutic options for the optimal management of MS.

Moreover, there is still lack of evidence-based approaches to incorporate patients' preferences such as medication disutility into the shared decision making process [29]. As our understanding of MS risk continues to be refined, how to account for the uncertain risks, benefits, and preferences at the individual level is a current challenge for the practice of personalized medicine.

## The proposed solution: bringing together advances in MS treatment and Neuroeconomics

The expected utility theory states that decision makers choose between risky or uncertain options by comparing their expected utility values (i.e.,: the weighted sums obtained by adding the utility values of outcomes multiplied by their respective probabilities) [30]. More importantly, patients' preferences and physicians' recommendations will change depending on the utility function of their current health status. For example, patients at low risk of developing MS progression may prefer to avoid 'risky' treatments (as they have low gains while having a risk of developing side effects), whereas high-risk patients would prefer the most effective treatment even if need to take higher risks (as they have a higher chance of having a progression leading to more disability) (Fig. 2) [24, 31].

## Rationale

Neuroeconomics is the science that studies the principles of how we make decisions [30, 32]. The neuroscience of decision making is based on behavioral economic concepts and mathematical approaches, such as game theory, to predict and model how people make their own choices [33].

The application of principles from Neuroeconomics (decision neuroscience) will facilitate the recognition of physicians' therapeutic preferences and beliefs about DMT for MS in the real world [34] (Fig. 3). Given the greater availability of treatment options, MS treatment will likely become more challenging. It requires a fine balance between the modest benefits of the less expensive, safer, and traditional DMTs versus new agents, usually more costly with potential harmful side effects. The so called 'intermediate agents' (e.g., Fingolimod) may have a 'decoy effect' (Phenomenon whereby consumers tend to have a specific change in preference between two options when also presented with a third -less preferable- option becomes available) [35, 36].

There is limited evidence of the application of the expected utility theory to clinical scenarios from the physicians' perspective. A better understanding of physicians' beliefs and preferences under uncertainty would likely improve the quality of care, patients' satisfaction, and likely improve clinical outcomes by increasing awareness on the therapeutic inertia in MS.

## Objectives

1) To evaluate tolerance to risk and ambiguity among MS experts under situations of uncertainty.
2) To assess the prevalence of 'therapeutic inertia'.
3) To determine the influence of tolerance to risk/ambiguity, overconfidence and herding on medical decisions.

## Research questions

1) How MS experts' perceptions of risk and tolerance to ambiguity influence their recommendations?

**How Absolute Risk-Aversion Changes with Wealth**

| Type of Risk-Aversion | Description |
|---|---|
| Increasing absolute risk-aversion | As wealth increases, hold fewer dollars in risky assets |
| Constant absolute risk-aversion | As wealth increases, hold the same dollar amount in risky assets |
| Decreasing absolute risk-aversion | As wealth increases, hold more dollars in risky assets |

**How Absolute Risk-Aversion Changes with Health**

| Type of Risk-Aversion | Description |
|---|---|
| Increasing absolute risk-aversion | As **health** increases (healthy), less interested in **risky treatments** |
| Constant absolute risk-aversion | As health is stable, hold the same interest in **risky treatments** |
| Decreasing absolute risk-aversion | As **health** decreases (sicker), more interested in **risky treatments** |

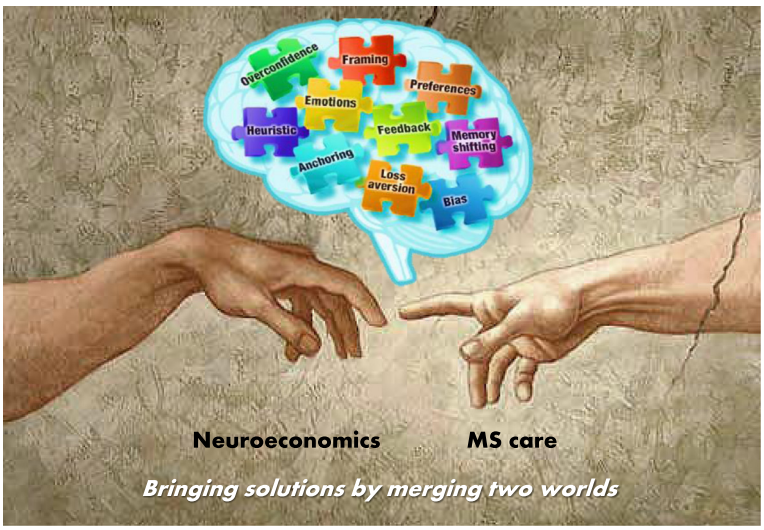
**Fig. 2** Illustrative comparison of risk aversion changes as a function of wealth and health

2) What is the prevalence of therapeutic inertia among physicians with expertise in MS?
3) What is the impact of tolerance to risk and ambiguity, overconfidence and herding on therapeutic decisions?

## Methods

We are proposing a prospective web-based study comprising 20 MS case-vignettes, 3 standardized surveys, and 4 behavioral experiments (see Additional file 1).

MS case-scenarios were derived from the most common situations in clinical practice as identified by experts in the field. Surveys include three standardized questions related to aversion to risk from The German Socio-Economic Panel (SOEP) study. The SOEP is a longitudinal study of private households that include household composition, occupational biographies, employment, earnings, health and satisfaction indicators [37, 38]. The English version is available online [39].




NeuroEconSolutions© by Gustavo Saposnik

**Fig. 3** Framework

Based on our previous work including a systematic literature review of studies evaluating cognitive biases or distortions in the medical field, we selected tolerance to risk and ambiguity, overconfidence, herding, and decisions about someone else [40]. We used the physician's reaction to uncertainty test to assess tolerance to risk or ambiguity in patient care [41]. This questionnaire comprised an initial pool of 61 items [41]. Factor analysis of the 428 respondents revealed a high accuracy (Cronbach's alpha = 0.90). The short version of this questionnaire includes five questions [42].

Behavioral experiments were designed to elicit risk and ambiguity aversion in the health and financial domains [43, 44], herding (decisions influenced by other colleagues) [45], decisions about someone else vs. own, and overconfidence (perception that own judgments are more accurate or in the top 50 % of the studied population) [46].

### Participants

Neurologists actively involved in the care of patients with MS from across Spain will be invited to participate in our study. Invitations are facilitated by the Spanish Society of Neurology (Sociedad Española de Neurologia). We use Qualtrics platform for the design and implementation of our study. It is expected physicians will be able to complete the study within 30 min.

Participating physicians will receive fair market compensation for the time involved in completing the survey.

### Outcome measures

The primary outcome of the study is the proportion of participants who exhibit aversion to ambiguity and therapeutic inertia [19, 43]. Ambiguity aversion is defined as is a preference for known risks over unknown risks [43]. This can be elicited through the experiments #16 and #17 in the health and financial domains (Additional file 1).

Therapeutic inertia will be assessed based on the selected treatment options in case-scenarios with recurrent relapses, appearance of new brain lesions in follow up MRI's while taking a disease modifying agent over a specified period. Secondary outcome measures include the association between risk aversion, overconfidence, and herding with therapeutic decisions and the assessment of therapeutic inertia using different criteria.

#### *Sample size calculation*

Based on the results of pilot studies evaluating other medical conditions (e.g., atrial fibrillation) and our systematic review on the frequency of cognitive distortions affecting physicians [40], we require a sample size of 120 physicians (60 per group) (Table 3) to reach 90 % power to detect a conservative 20 % absolute difference in therapeutic inertia between participants exposed and not exposed to cognitive distortions.

**Table 3** Sample size calculation

| Power[a] | 90 % | 85 % | 80 % |
|---|---|---|---|
| N (per group) | 60 | 53 | 46 |

[a]The power was calculated to detect a 20 % absolute difference between groups (40 % vs 20 %) with an alpha of 5 % (two-sided) for all of the calculations in the table

#### *Feasibility*

the study interventions are simple and doable. The protocol includes clinical scenarios commonly observed in clinical practice. According to the Spanish Neurological Society (Sociedad Española de Neurología-SEN), there are over 1600 neurologists, 13 specialized MS centers comprising approximately 200 specialists in the field in Spain. Assuming a low response rate of 50 %, the completion of our study is feasible considering the required sample size to reach a power of 90 % with an alpha of 5 %.

### Analytical plan

To address objective 1, we will characterize participants' risk and ambiguity aversion as identified by the corresponding experiments (Additional file 1, behavioral battery questions (Q) #1 to 4.

To address objective 2, we evaluate therapeutic inertia (TI) as elicited by 10 case-vignettes. We will created a TI score representing the number of cases that participants did not escalate treatment (numerator) over 10 (denominator) multiplied by 100. The diversity of case-scenarios will also allow the analysis of therapeutic inertia using different criteria (e.g., modified Rio score, European Medicines Agency, isolated GAD-enhanced lesions).

To address objective 3, we will complete a univariate and multivariable analysis to determine the influence of risk aversion, tolerance to ambiguity, overconfidence and herding on therapeutic decisions and TI score.

Chi squared tests will be used to compare categorical variables; *t*-test or Kruskal-Wallis tests will be used to compare mean and median differences for continuous variables. The primary analysis will evaluate the association between physicians' responses in the behavioral component of the survey with responses in the case-scenarios. A multivariable analysis will be completed to determine the association between physicians' characteristics with the primary outcome of interest. Adjustment includes the following variables: age, sex, years of experience, expertise, volume of MS patients seen per week, and practice setting (academic vs. community). All tests were 2-tailed, and p-values $<0.05$ will be considered significant.

#### *Knowledge translation strategies*

We plan to take a multifaceted approach to knowledge translation, targeting the following audiences for communication: 1) Neurologist, 2) the clinical academic

community, 3) the media, 4) policy-makers, and 5) MS patients and their families. We expect to generate high impact publications and media interest to inform the public and influence MS care programs. This work is also expected to increase awareness about therapeutic inertia among MS experts and to contribute toward new guidelines for the management of MS. We are working with key stakeholders to discuss the most effective dissemination strategy and target the key messages for all audiences.

## Discussion

Patients and physicians caring for patients with MS are confronted with important uncertainties concerning the diagnosis, prognosis, disease course, and disease-modifying therapies. In the recent years, new therapeutic alternatives became available for management of MS [5, 47]. These advances were achieved by targeting different pathophysiological mechanisms, producing more effective DMTs, but accompanied by either higher risk of infections, or more serious side effects [48]. As a result, MS experts have an expanded therapeutic arsenal compared to a decade ago. Decisions are not merely about the selection of an injectable interferon or Glatiramer (given daily or every other day) usually accompanied by skin reactions or flu-like symptoms, but rather the individual selection of the most appropriate DMT (e.g., dose, administration type, efficacy and safety profile) according to disease severity, patient's clinical status and preference. Consequently, more effective agents are now more accessible for MS patients who failed traditional DMT [5, 49].

Interestingly, physicians have limited education in both risk management and in formal training in decision making [50, 51].

We are proposing a novel approach in expanding research of MS care by combining case-vignettes with the assessment of cognitive distortions through experiments in Neuroeconomics (Decision Neuroscience). The application of Neuroeconomics' principles may help overcome those barriers by identifying and increasing awareness about cognitive distortions (e.g., overconfidence, tolerance to risk and ambiguity, etc.) that may lead to suboptimal decisions (e.g., therapeutic inertia) [18, 25, 52].

This study will provide evidence about: i) how MS experts make decisions under uncertainty, ii) how MS experts would change their preferences based on their tolerance to risk and ambiguity, iii) the prevalence of therapeutic inertia based on different criteria for escalating therapy (modified Rio, European Medicines Agency), and iv) the influence of cognitive distortions on therapeutic inertia.

DIScUTIR MS is designed as a pilot study to determine the feasibility of assessing tolerance to risk and ambiguity, therapeutic inertia, and associated factors among practicing physicians with expertise in MS.

The results of our study will also facilitate crucial information to understand current MS care practices and how physicians' preferences (e.g., risk aversion) have a global impact on medical and daily life decisions.

Some limitations need to be acknowledged. First, the small sample size of MS experts from a single country (Spain) would limit the generalizability of the results. However, DIScUTIR MS is designed as a pilot study to determine the feasibility of a larger worldwide study. Second, the concept and definition of therapeutic inertia applied to MS care is not widely disseminated. Some colleagues may also argue about the absence of an accepted definition of therapeutic inertia in MS care. However, we used a widely acceptable definition of TI supported by studies showing health care improvements in the management of key and widely prevalent conditions (i.e.,: blood pressure and diabetes).

Despite the aforementioned limitations, our study will increase physicians' awareness of crucial situations under uncertainty in the management of MS. The results of DIScUTIR MS will provide a starting point to ignite discussions about a widely accepted definition of therapeutic inertia in MS care. This is relevant considering the lack of MS guidelines concerning clinical scenarios under uncertainty or progression of disease [53, 54].

The identification of clinical or radiological progression in MS should at least set the time of 'therapeutic momentum' to consider escalating treatment, especially when cost-effective options are available. In this setting, physicians may want to take that opportunity to discuss risk-benefit scenarios in a way similar to how financial advisors assess their clients' preferences and risk tolerance when advising about a variety of investment portfolios. An open discussion in risky situations following the appropriate documentation of disease progression would ameliorate the therapeutic inertia and may lead to more optimal decisions in the care of patients with MS.

### Ethics approval

The protocol was approved by the Research Ethics Board of St. Michael's Hospital, University of Toronto. Consent will be obtained by agreeing to participate in the study.

### Availability of data and materials

The appendix contains all details of the protocol.

## Additional file

**Additional file 1:** Case-vignettes and behavioral experiments. (DOCX 601 kb)

### Abbreviations

DMT: disease modifying agents; EMA: European Medicines Agency; MS: multiple sclerosis; SOEP: Socio-Economic Panel Study.

### Competing interests

Drs. Maurino and Prefasi are employees in the Medical Department of Roche Pharma. They do not hold any stocks or shares in Roche Pharma that may in any way gain or lose financially from the publication of this manuscript. Drs. Saposnik, Perez Sempere, Selchen and Raptis have no financial competing interest.

### Authors' contributions

We declare that we have participated in the (conception, design of the study, drafting the manuscript and made a critical revision of the manuscript). Dr. Saposnik was responsible for obtaining funds. Dr. Saposnik is supported by the Clinician-Scientist Award from Heart and Stroke Foundation Canada (HSFC). All authors read and approved the final manuscript.

### Acknowledgements

The study was sponsored by the Sociedad Española de Neurologia (SEN) and funded by an operating grant from Roche Farma Spain.
The authors are most grateful to all physicians participating in the study.
We are indebted to Professors Christian Ruff and Philippe Tobler from the Department of Economics at the University of Zurich for their guidance when designing the behavioral experiments and survey selection.
The sponsors were not involved in the design, execution, analysis, and interpretation or reporting of the results.
We would like to thank the Li Ka Shing Knowledge Institute at St Michael's Hospital for the organization of the study.

### Funding

The study is supported by Roche Pharma and the Spanish Society of Neurology (SEN).

### Author details

[1]Division of Neurology, Department of Medicine, St. Michael's Hospital, University of Toronto, 55 Queen St E, Toronto, ON M5C 1R6, Canada. [2]Neuroeconomics and Decision Neuroscience, Department of Economics, University of Zurich, Zurich, Switzerland. [3]Department of Neurology, Hospital General Universitario de Alicante, Alicante, Spain. [4]Applied Health Research Center, Li Ka Shing Knowledge Institute, St. Michael's Hospital, University of Toronto, Toronto, Canada. [5]Neuroscience Area, Medical Department, Roche Farma, Madrid, Spain.

Received: 24 December 2015 Accepted: 21 April 2016
Published online: 04 May 2016

### References

1. English C, Aloi JJ. New fda-approved disease-modifying therapies for multiple sclerosis. Clin Ther. 2015;37:691–715.
2. Bruck W, Gold R, Lund BT, Oreja-Guevara C, Prat A, Spencer CM, et al. Therapeutic decisions in multiple sclerosis: moving beyond efficacy. JAMA Neurol. 2013;70:1315–24.
3. Sormani MP, Bruzzi P. Can we measure long-term treatment effects in multiple sclerosis? Nat Rev Neurol. 2015;11:176–82.
4. D'Amico E, Leone C, Caserta C, Patti F. Oral drugs in multiple sclerosis therapy: an overview and a critical appraisal. Expert Rev Neurother. 2015;15:803–24.
5. Feinstein A, Freeman J, Lo AC. Treatment of progressive multiple sclerosis: What works, what does not, and what is needed. Lancet Neurol. 2015;14:194–207.
6. Hartung HP, Aktas O, Boyko AN. Alemtuzumab: a new therapy for active relapsing-remitting multiple sclerosis. Mult Scler. 2015;21:22–34.
7. Sempere AP, Gimenez-Martinez J. Safety considerations when choosing the appropriate treatment for patients with multiple sclerosis. Expert Opin Drug Saf. 2014;13:1287–9.
8. Ontaneda D, Cohn S, Fox RJ. Risk stratification and mitigation in multiple sclerosis. Mult Scler Relat Disord. 2014;3:639–49.
9. Ransohoff RM, Hafler DA, Lucchinetti CF. Multiple sclerosis-a quiet revolution. Nat Rev Neurol. 2015;11:134–42.
10. Sormani MP, De Stefano N. Defining and scoring response to ifn-beta in multiple sclerosis. Nat Rev Neurol. 2013;9:504–12.
11. Sormani MP, Rio J, Tintore M, Signori A, Li D, Cornelisse P, et al. Scoring treatment response in patients with relapsing multiple sclerosis. Mult Scler. 2013;19:605–12.
12. Freedman MS, Forrestal FG. Canadian treatment optimization recommendations (tor) as a predictor of disease breakthrough in patients with multiple sclerosis treated with interferon beta-1a: analysis of the prisms study. Mult Scler. 2008;14:1234–41.
13. Freedman MS, Selchen D, Arnold DL, Prat A, Banwell B, Yeung M, et al. Treatment optimization in ms: Canadian ms working group updated recommendations. Can J Neurol Sci. 2013;40:307–23.
14. Bermel RA, You X, Foulds P, Hyde R, Simon JH, Fisher E, et al. Predictors of long-term outcome in multiple sclerosis patients treated with interferon beta. Ann Neurol. 2013;73:95–103.
15. Prosperini L, Mancinelli CR, De Giglio L, De Angelis F, Barletta V, Pozzilli C. Interferon beta failure predicted by ema criteria or isolated mri activity in multiple sclerosis. Mult Scler. 2014;20:566–76.
16. O'Connor PJ, Sperl-Hillen JAM, Johnson PE, Rush WA, Biltz G. Clinical inertia and outpatient medical errors. In: Henriksen K, Battles JB, Marks ES, Lewin DI, editors. Advances in patient safety: from research to implementation (volume 2: Concepts and methodology). Rockville: Agency for Healthcare Research and Quality (US); 2005.
17. Mohan AV, Phillips LS. Clinical inertia and uncertainty in medicine. JAMA. 2011;306:383. author reply 383–384.
18. Phillips LS, Branch WT, Cook CB, Doyle JP, El-Kebbi IM, Gallina DL, et al. Clinical inertia. Ann Intern Med. 2001;135:825–34.
19. Okonofua EC, Simpson KN, Jesri A, Rehman SU, Durkalski VL, Egan BM. Therapeutic inertia is an impediment to achieving the healthy people 2010 blood pressure control goals. Hypertension. 2006;47:345–51.
20. Huang LY, Shau WY, Yeh HL, Chen TT, Hsieh JY, Su S, et al. A model measuring therapeutic inertia and the associated factors among diabetes patients: A nationwide population-based study in taiwan. J Clin Pharmacol. 2015;55:17–24.
21. Escobar C, Barrios V, Alonso-Moreno FJ, Llisterri JL, Rodriguez-Roca GC, Prieto MA, et al. New blood pressure control goals, more rational but facilitating therapeutic inertia? J Hypertens. 2013;31:2462.
22. Turner BJ, Hollenbeak CS, Weiner M, Ten Have T, Tang SS. Effect of unrelated comorbid conditions on hypertension management. Ann Intern Med. 2008;148:578–86.
23. Kerr EA, Zikmund-Fisher BJ, Klamerus ML, Subramanian U, Hogan MM, Hofer TP. The role of clinical uncertainty in treatment decisions for diabetic patients with uncontrolled blood pressure. Ann Intern Med. 2008;148:717–27.
24. Glimcher P, Fehr E. Neuroeconomics: decision making and the brain. San Diego: Academic; 2014.
25. Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. Acad Med. 2003;78:775–80.
26. Elstein AS, Schwartz A. Clinical problem solving and diagnostic decision making: selective review of the cognitive literature. BMJ. 2002;324:729–32.
27. Reach G. Clinical inertia, uncertainty and individualized guidelines. Diabetes Metab. 2014;40:241–5.
28. Ye S. Medical decision making and the counting of uncertainty. Circulation. 2014;129:2500–2.
29. Fontana M, Asaria P, Moraldo M, Finegold J, Hassanally K, Manisty CH, et al. Patient-accessible tool for shared decision making in cardiovascular primary prevention: balancing longevity benefits against medication disutility. Circulation. 2014;129:2539–46.
30. Platt ML, Huettel SA. Risky business: the neuroeconomics of decision making under uncertainty. Nat Neurosci. 2008;11:398–403.
31. Tur C, Tintore M, Vidal-Jordana A, Bichuetti D, Nieto Gonzalez P, Arevalo MJ, et al. Risk acceptance in multiple sclerosis patients on natalizumab treatment. PLoS One. 2013;8:e82796.
32. Lee D. Neuroeconomics: Best to go with what you know? Nature. 2006;441:822–3.
33. d'Acremont M, Bossaerts P. Decision making: How the brain weighs the evidence. Curr Biol. 2012;22:R808–10.

34. Saposnik G, Johnston SC. Decision making in acute stroke care: learning from neuroeconomics, neuromarketing, and poker players. Stroke. 2014;45:2144–50.
35. Ariely D. Predictably irrational: the hidden forces that shape our decisions. New York: HarperCollins Publishers; 2008.
36. Hu J, Yu R. The neural correlates of the decoy effect in decisions. Front Behav Neurosci. 2014;8:271.
37. Wagner GG, Frick JR, Schupp J, Panel DIfWPDS-Ö. The German socio-economic panel study (soep): scope, evolution and enhancements. Berlin: DIW Berlin; 2007.
38. Dohmen T, Falk A, Huffman D, Sunde U, Schupp J, Wagner GG. Individual risk attitudes: measurement, determinants, and behavioral consequences. 2011.
39. Panel DIfWPDS-Ö, Sozialforschung TI. Soep 2014 – erhebungsinstrumente 2014 (welle 31) des sozio-oekonomischen panels: Personenfragebogen, altstichproben. Berlin: DIW Berlin / SOEP; 2014.
40. Saposnik G, Redelmeier D, Ruff CC, Tobler PN. Cognitive distortions and associated with medical desicions: a systematic review. BMC Med. 2015.
41. Gerrity MS, DeVellis RF, Earp JA. Physicians' reactions to uncertainty in patient care. A new measure and new insights. Med Care. 1990;28:724–36.
42. Gerrity M, White K, DeVellis R, Dittus R. Physicians' reactions to uncertainty: refining the constructs and scales. Motiv Emot. 1995;19:175–91.
43. Levy I, Snell J, Nelson AJ, Rustichini A, Glimcher PW. Neural representation of subjective value under risk and ambiguity. J Neurophysiol. 2010;103:1036–47.
44. Anderson LR, Mellor JM. Predicting health behaviors with an experimental measure of risk preference. J Health Econ. 2008;27:1260–74.
45. David E, Jon K. Information cascades (chapter 16). Cambridge: Cambridge University Press; 2010.
46. Cesarini D, Sandewall Ö, Johannesson M. Confidence interval estimation tasks and the economics of overconfidence. J Econ Behav Organ. 2006;61:453–70.
47. Ontaneda D, Fox RJ, Chataway J. Clinical trials in progressive multiple sclerosis: lessons learned and future perspectives. Lancet Neurol. 2015;14:208–23.
48. Mahad DH, Trapp BD, Lassmann H. Pathological mechanisms in progressive multiple sclerosis. Lancet Neurol. 2015;14:183–93.
49. Kopke S, Solari A, Khan F, Heesen C, Giordano A. Information provision for people with multiple sclerosis. Cochrane Database Syst Rev. 2014;4: CD008757.
50. Studdert DM, Mello MM, Sage WM, DesRoches CM, Peugh J, Zapert K, et al. Defensive medicine among high-risk specialist physicians in a volatile malpractice environment. JAMA. 2005;293:2609–17.
51. Dijkstra IS, Pols J, Remmelts P, Brand PL. Preparedness for practice: a systematic cross-specialty evaluation of the alignment between postgraduate medical education and independent practice. Med Teach. 2015;37:153–61.
52. Saposnik G, Redelmeier D, Ruff CC, Tobler PN. Cognitive distortions and associated with medical desicions: a systematic review Med Decis Mak. 2015.
53. Kalincik T, Cutter G, Spelman T, Jokubaitis V, Havrdova E, Horakova D, et al. Defining reliable disability outcomes in multiple sclerosis. Brain. 2015;138: 3287–98.
54. Lublin FD, Reingold SC, Cohen JA, Cutter GR, Sorensen PS, Thompson AJ, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. Neurology. 2014;83:278–86.

# B. Appendix to Study 2: ii) Results

frontiers in Neurology

ORIGINAL RESEARCH
published: 01 March 2017
doi: 10.3389/fneur.2017.00065

Check for updates

## Decision-making in Multiple Sclerosis: The Role of Aversion to Ambiguity for Therapeutic Inertia among Neurologists (DIScUTIR MS)

*Gustavo Saposnik[1,2,3]*, Angel P. Sempere[4], Daniel Prefasi[5], Daniel Selchen[1], Christian C. Ruff[2], Jorge Maurino[5†] and Philippe N. Tobler[2†]*

[1]*Division of Neurology, Stroke Outcomes and Decision Neuroscience Research Unit, Department of Medicine, St. Michael's Hospital, University of Toronto, Toronto, ON, Canada,* [2]*Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Zurich, Switzerland,* [3]*Li Ka Shing Knowledge Institute, St. Michael's Hospital, University of Toronto, Toronto, ON, Canada,* [4]*Department of Neurology, Hospital General Universitario de Alicante, Alicante, Spain,* [5]*Neuroscience Area, Medical Department, Roche Farma, Madrid, Spain*

OPEN ACCESS

**Edited by:**
*Björn Tackenberg, Philipps University of Marburg, Germany*

**Reviewed by:**
*Samar S. Ayache,*

**Objectives:** Limited information is available on physician-related factors influencing therapeutic inertia (TI) in multiple sclerosis (MS). Our aim was to evaluate whether physicians' risk preferences are associated with TI in MS care, by applying concepts from behavioral economics.

This paper is published as Saposnik G, Sempere AP, Raptis R, Prefasi D, Selchen D, Maurino J, Decision-making in Multiple Sclerosis: The Role of Aversion to Ambiguity for Therapeutic Inertia among Neurologists (DIScUTIR MS). Front Neurol., 8, 65, 2017

frontiers in Neurology

ORIGINAL RESEARCH
published: 01 March 2017
doi: 10.3389/fneur.2017.00065

# Decision-making in Multiple Sclerosis: The Role of Aversion to Ambiguity for Therapeutic Inertia among Neurologists (DIScUTIR MS)

*Gustavo Saposnik[1,2,3]*, Angel P. Sempere[4], Daniel Prefasi[5], Daniel Selchen[1], Christian C. Ruff[2], Jorge Maurino[5†] and Philippe N. Tobler[2†]*

[1]*Division of Neurology, Stroke Outcomes and Decision Neuroscience Research Unit, Department of Medicine, St. Michael's Hospital, University of Toronto, Toronto, ON, Canada,* [2]*Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Zurich, Switzerland,* [3]*Li Ka Shing Knowledge Institute, St. Michael's Hospital, University of Toronto, Toronto, ON, Canada,* [4]*Department of Neurology, Hospital General Universitario de Alicante, Alicante, Spain,* [5]*Neuroscience Area, Medical Department, Roche Farma, Madrid, Spain*

OPEN ACCESS

***Edited by:***
*Björn Tackenberg, Philipps University of Marburg, Germany*

***Reviewed by:***
*Samar S. Ayache, Paris Est University, France*
*Christoph Alexander Mayer, University Hospital Frankfurt, Germany*

****Correspondence:***
*Gustavo Saposnik*
*saposnikg@smh.ca*

*[†]These authors have contributed equally to this work.*

***Specialty section:***
*This article was submitted to Multiple Sclerosis and Neuroimmunology, a section of the journal Frontiers in Neurology*

***Received:*** *21 October 2016*
***Accepted:*** *13 February 2017*
***Published:*** *01 March 2017*

***Citation:***
*Saposnik G, Sempere AP, Prefasi D, Selchen D, Ruff CC, Maurino J and Tobler PN (2017) Decision-making in Multiple Sclerosis: The Role of Aversion to Ambiguity for Therapeutic Inertia among Neurologists (DIScUTIR MS). Front. Neurol. 8:65. doi: 10.3389/fneur.2017.00065*

**Objectives:** Limited information is available on physician-related factors influencing therapeutic inertia (TI) in multiple sclerosis (MS). Our aim was to evaluate whether physicians' risk preferences are associated with TI in MS care, by applying concepts from behavioral economics.

**Design:** In this cross-sectional study, participants answered questions regarding the management of 20 MS case scenarios, completed 3 surveys, and 4 experimental paradigms based on behavioral economics. Surveys and experiments included standardized measures of aversion ambiguity in financial and health domains, physicians' reactions to uncertainty in patient care, and questions related to risk preferences in different domains. The primary outcome was TI when physicians faced a need for escalating therapy based on clinical (new relapse) and magnetic resonance imaging activity while patients were on a disease-modifying agent.

**Results:** Of 161 neurologists who were invited to participate in the project, 136 cooperated with the study (cooperation rate 84.5%) and 96 completed the survey (response rate: 60%). TI was present in 68.8% of participants. Similar results were observed for definitions of TI based on modified Rio or clinical progression. Aversion to ambiguity was associated with higher prevalence of TI (86.4% with high aversion to ambiguity vs. 63.5% with lower or no aversion to ambiguity; $p = 0.042$). In multivariate analyses, high aversion to ambiguity was the strongest predictor of TI (OR 7.39; 95%CI 1.40–38.9), followed by low tolerance to uncertainty (OR 3.47; 95%CI 1.18–10.2).

**Conclusion:** TI is a common phenomenon affecting nearly 7 out of 10 physicians caring for MS patients. Higher prevalence of TI was associated with physician's strong aversion to ambiguity and low tolerance of uncertainty.

**Keywords: multiple sclerosis, disease-modifying therapy, neuroeconomics, decision-making, risk aversion**

## INTRODUCTION

Making decisions in medical care is a complex task (1). Physicians have limited education in both risk management and decision-making at medical schools (2). Decisions based on erroneous assessments may result in incorrect patient and family expectations, and potentially suboptimal advice, treatment, and outcome.

In behavioral economics, *uncertainty* is a generic term that comprises risk and ambiguity. *Risk* applies to events with known probability (3). In contrast, *ambiguity* is a term reserved for events for which probabilities are unknown (3). Typically, people are averse to both ambiguity and risk, and the two aversions are independent of each other (4). Uncertainty is one of the most important contributing factors affecting decisions in medical care (5, 6). However, limited information is available regarding the role of aversion to ambiguity in medical decisions.

Appropriate multiple sclerosis (MS) care involves complex medical decisions as it requires consideration of multiple short- and long-term factors (e.g., imaging results, disease progression, patient's characteristics, and their preferences, etc.). No evidence of disease activity is emerging as a new standard for treatment response and may be associated with improved long-term disability outcomes. A more proactive management strategy, including earlier use of high-efficacy DMTs and close monitoring of the clinical and radiological response to treatment, is recommended to slow the progression of physical and cognitive impairments in patients with relapsing-remitting multiple sclerosis (RRMS) (7–9). Treatment escalation has been shown to reduce relapse rates, disability progression, and magnetic resonance imaging (MRI) activity (10).

Therapeutic inertia (TI) is a term introduced in 2006 to define the absence of treatment initiation or intensification in patients when treatment goals are unmet (11–14). In the context of MS, TI is defined as the lack of treatment initiation or escalation when there is evidence of disease activity (based on the clinical course and neuroimaging markers) (15, 16). It is possible that aversion to ambiguity contributes to TI as the probabilities of benefits with treatment escalation are typically less well known than with treatment continuation. To address this possibility, we need a better understanding of physician-related factors influencing decisions about DMTs and the prevalence of TI in MS care. The application of experiments from behavioral economics would facilitate the recognition of physicians' therapeutic preferences and beliefs about DMTs for MS in the real world (17).

We hypothesized that physicians' ambiguity aversion or low tolerance to uncertainty are associated with TI and clinical decisions in MS care. In the present study, we thus assessed the prevalence of TI (and associated contributing factors) in typical clinical decisions among physicians caring for MS patients across Spain.

## MATERIALS AND METHODS

We conducted a web-based study using the Qualtrics platform (http://qualtrics.com). The study comprised 20 MS case-vignettes, 3 standardized surveys, and 4 behavioral experiments among practicing neurologists from Spain from November 3, 2015 to March 31, 2016 (see protocol published elsewhere) (15). In brief, participants answered three components in the following order: (i) demographic information, (ii) behavioral experiments/surveys, and (iii) case scenarios. Responses from case scenarios were analyzed in light of responses from the behavioral component. MS case scenarios were derived from the most common situations in clinical practice as identified by experts in the field (Drs. Daniel Selchen and Angel P. Sempere). Behavioral experiments were designed to assess risk and ambiguity aversion in the health and financial domains (exposure) (**Figure 1**) (15, 18, 19). Ambiguity aversion is defined as dislike for events with unknown probability over events with known probability (18). For example, an ambiguity-averse individual would rather choose a treatment where the probability of benefits or side effects are known (even if these are somewhat unfavorable) over one where these probabilities are unknown. Specifically, participants were asked to choose between a visual option with known 50/50 probability of winning €400 or €0 and an option with unknown probability of the same outcomes. Gray bars represented the degree to which the winning probability was unknown (**Figure 1**). The degree of ambiguity aversion was defined as the proportion of times participants chose the 50/50 option over the ambiguous option. As the overall level of ambiguity aversion was pronounced in our sample (mean 61.7% preference for 50/50 option, i.e., the option with known probabilities) and to avoid using an arbitrary criterion, we classified participants as highly ambiguity averse if they chose the 50/50 (known probability) option in each of the nine scenarios (**Figure 1**) (20). In order to evaluate the consistency of the relationship with the primary outcome, we also analyzed another definition of ambiguity aversion (choice of the known probability option instead of the option with the 50% unknown probability in scenario 5; **Figure 1**).

In principle, risk aversion is another factor that may influence clinical decisions (21). Risk aversion is defined as the tendency to prefer safe payoffs over probabilistic payoffs when the expected value of both options is identical (4, 18). A risk-averse patient would thus prefer a treatment that provides a small improvement with certainty over a treatment that provides a larger or no improvement with equal chance (50/50). We evaluated risk aversion by identifying the safe amount for which a participant was indifferent between the safe and the risky option (22). This indifference amount, called certainty equivalent, reflects the value associated to the risky option and facilitates comparison between participants. For example, participants were asked what would be the minimal certain payoff that they would prefer over the equiprobable gamble of winning €400 or €0 (expected value of €200). The degree of risk aversion of each individual corresponded to the difference of the expected value of the risky option (€200) minus the participant's response (proxy of certainty equivalent). A similar visual design and methodology was used to elicit aversion to risk and ambiguity in the health domains (questions #15 and #17) (15). Participants were asked to choose between Treatment A (50% probability of survival) and "Treatment B" (the probability of survival is unknown) with the gray bars quantifying how much is unknown about the probability of survival.

We also used two standardized surveys to assess physicians' willingness to take risks and tolerance to uncertainty. The German

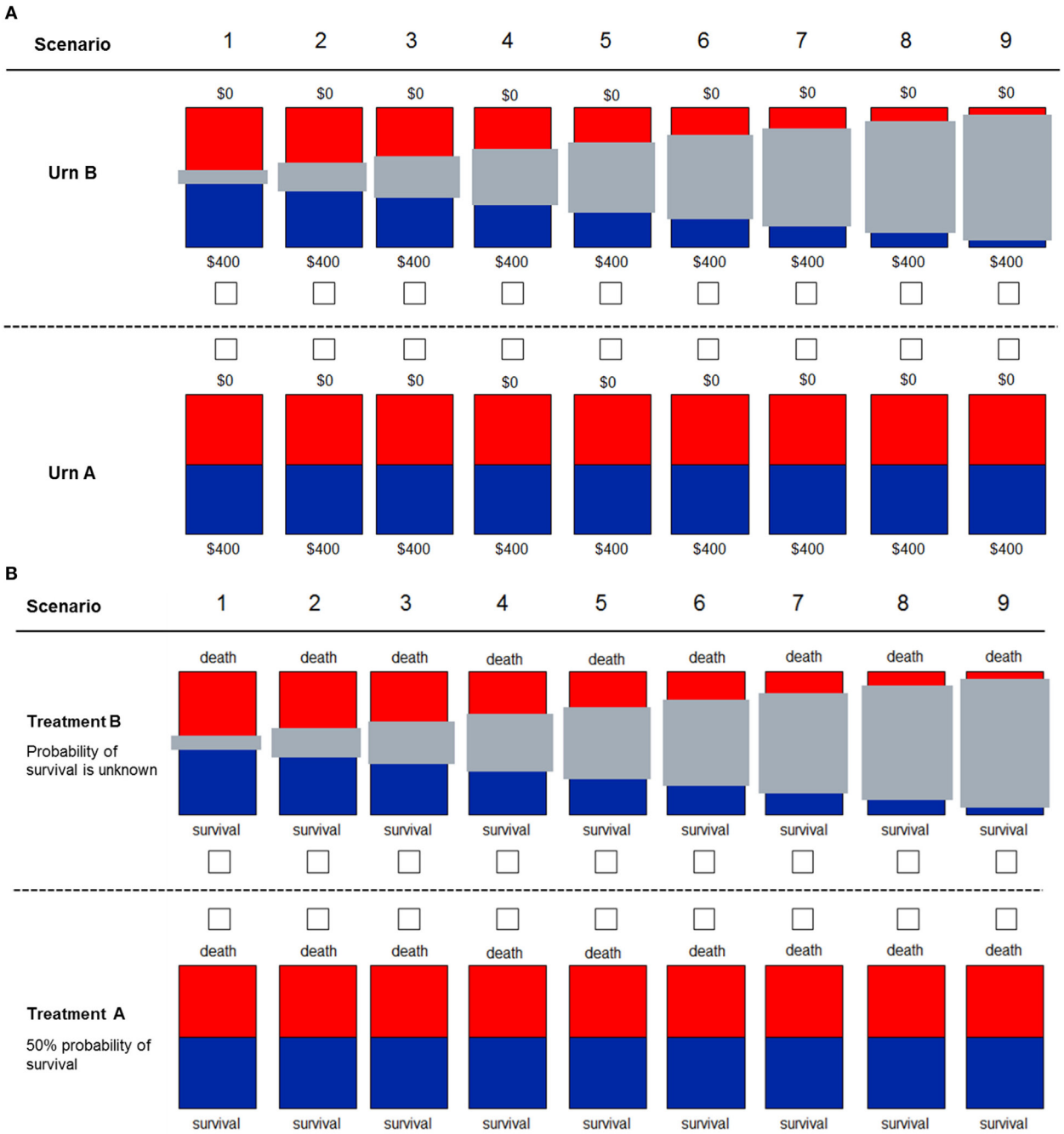



**FIGURE 1 | Decision scenarios used to measure ambiguity in financial (A) and health (B) domains**. Participants were told to imagine two different types or urns. For urn type A, they knew that 50% of the balls were red and the other 50% were blue. For urn type B, they did not know the exact proportion of blue to red balls, with the gray bar representing the unknown proportion of balls. For the financial domain, participants knew that if they drew a blue ball, they would win the full amount of $400. If they drew a red ball, they would win $0. For the health domain, participants decided between two treatments for a patient. With "Treatment A," the patient had a 50% probability of survival. With "Treatment B," the exact probability of survival was unknown, with the gray bar representing the unknown probability.

Socio-Economic Panel (SOEP) is a validated survey that evaluates willingness to take risks in different domains (financial matters, own health, driving, own occupation, etc.) (23). We used questions of the form: "How would you rate your willingness to take risks in the following areas…."? Areas included financial matters, driving, occupation, etc., and responses could range from 0 (not at all) to 10 (very much).

The second survey measured physicians' tolerance to uncertainty in patient care, using the reaction to uncertainty test (24). It comprises five questions to be rated from 0 to 5 that when added

gives a total score (25). Low tolerance to uncertainty was defined as values below the median of the total score. Further details of the protocol were published elsewhere (15).

## Participants

Practicing neurologists actively involved in the care of patients with MS from across Spain were invited to participate in our study by the Spanish Society of Neurology (Sociedad Española d8e Neurologia-SEN). Physicians whose practice was primarily in caring for MS patients were classified as "MS specialists." All participants received compensation for completing the survey.

## Definitions

For the primary analysis, disease activity was defined as a clinical relapse plus the presence of new brain lesions in follow-up MRI scans with at least one gadolinium-enhancing lesion (26, 27). In a sensitivity analysis, we also used the European Medicines Agency (EMA) and the modified Rio criteria to evaluate variations in TI. For example, the EMA approves escalating therapy from interferon-beta to natalizumab or fingolimod in patients who had at least one relapse in the previous year and either ≥9 T2 hyperintense lesions or ≥1 gadolinium-enhancing T1 lesion on brain MRI (26, 27). The high-risk profile according to the modified Rio score includes either an MRI with more than 5 new T2 lesions (1 point) or 1 relapse in the first year (1 point) or two relapses within the first year of treatment (2 points) or the combination of these criteria (8, 28). The use of these definitions combining a clinical relapse and MRI activity is consistent with recent evidence regarding the risk of treatment failure among patients receiving interferon-beta (29).

Disease progression was defined as at least one point worsening from baseline in the Expanded Disability Status Scale (EDSS) score (Table S1 in Supplementary Material) (28).

Recent meta-analysis confirmed that alemtuzumab, natalizumab, and fingolimod are the best available choices for preventing clinical relapses in patients with RRMS (30). However, there is no consensus algorithm available despite the publication of national or regional recommendations (8, 16, 26, 31–33). As a result, the current landscape of DMTs for the treatment of RRMS includes first-line therapies (beta interferons, glatiramer acetate, teriflunomide, and dimethyl fumarate) and second-line therapies (fingolimod, natalizumab, and alemtuzumab). For the present analysis, we used the aforementioned scheme according to the current clinical practice.

## Outcome Measures

The primary outcome of the study was the proportion of participants who exhibited TI and its association with aversion to ambiguity (14, 18). TI (presence/absence) was determined as the lack of escalation of therapy given disease activity while patients received a DMT in at least one case scenario.

Secondary outcome measures included the association between tolerance to uncertainty, risk aversion, and the SOEP surveys, on the one hand, and with TI and therapeutic decisions, on the other hand.

## Statistical Analysis

The primary analysis assessed the possible association between physicians' aversion to ambiguity and TI. A multiple logistic regression analysis with backward selection was completed to determine the association between physicians' characteristics with the primary outcome of interest. We included the following explanatory variables: age, gender, MS patients seen per week, practice setting (academic vs. non-academic), % of time devoted to clinical care, coauthor in a peer-reviewed publication within the last 3 years (yes/no), attendance to the European Committee for Treatment and Research in Multiple Sclerosis 2015 annual meeting, risk aversion, overconfidence, tolerance to uncertainty (above/below the median), willingness to take risks in all domains (SOEP survey—above/below the median), and herding (following recommendations made by another colleague). As there was a high correlation between MS specialists (self-defined) and number of MS patients assessed per week (Spearman's rho = 0.58; $p < 0.001$), only the latter was entered in the multivariate analysis. Linear regression analysis was used to test for a relation between the number of patients assessed per week and the outcomes of interest. A sensitivity analysis was completed by using different criteria of TI and building models including all variables of interest (Supplementary Material).

All tests were 2-tailed, and $p$-values <0.05 were considered significant. We calculated the power of the study for the primary outcome of interest with an alpha error level of 0.05 and found that we had 100% power to detect a 27% difference between groups for the primary outcome measure.

The study was approved by the Research Ethics Board of St. Michael's Hospital, University of Toronto, ON, Canada.

# RESULTS

Out of the 161 neurologists who were invited to participate in the study from representative areas of Spain, 136 cooperated (cooperation rate 84.5%) and 96 completed the survey (response rate 60%). There was representation from all regional territories except the Canary Islands (Figure S1 in Supplementary Material).

Overall, the mean (SD) age was 39.5 (±8.5) years; 51 (53%) were female. Two-thirds primarily focused their practice on MS care ($n = 64$; 66.7%). The mean years in practice was 14, commonly assessing 20 (±15) MS patients per week. **Table 1** summarizes baseline characteristics of the study population.

For the measurement of risk preferences, the mean safe payoffs were €200 (±33) in the financial domain and 12.3 (±4.3) years in the health domain.

For the measurement of ambiguity, total aversion to unknown probability (all nine scenarios) was observed in 23% of participants in the financial domain and 27.1% for the health domain. For the scenario where the ambiguous option contained 50% unknown probability (scenario 5), 59.4% of participants chose the known probability (50/50) option in the financial domain and 73.7% in the health domain. The median time for completing the study was 39 min (IQR 30–52 min).

Therapeutic inertia was present in 68.8% of participants. Similar findings were observed when we applied the modified Rio criteria (61.5%), modified Rio or neurological progression (67.7%), but

**TABLE 1 | Baseline characteristics of participants.**

| Characteristics | |
|---|---|
| **Age** (mean ± SD), in years | 39.5 ± 8.5 |
| **Sex** | **No. of participants (%)** |
| Male | 45 (46.9) |
| Female | 51 (53.1) |
| **Specialty** | |
| Multiple sclerosis (MS) specialist | 64 (66.7) |
| General neurologist who care for MS patients | 32 (33.3) |
| **Practice setting** | |
| Academic | 48 (50.0) |
| Community | 26 (27.1) |
| Both (academic and community) | 21 (21.9) |
| Other | 1 (1.0) |
| **% time in clinical practice** | |
| >75% | 70 (72.9) |
| **Years in practice**, mean ± SD | 14.1 ± 10 |
| **MS patients seen per week**, mean ± SD | 20 ± 15 |
| **Attended latest European Committee for Treatment and Research in Multiple Sclerosis conference** | 56 (58) |
| **Author of a peer-reviewed publication in the last 3 years** | 79 (82.3) |

TI was less common (29.2%) when we applied the EMA criteria. TI was less common among MS specialists (**Table 2**). Moreover, a higher number of MS patients seen per week were associated with a significantly lower risk of TI. Linear regression analysis suggests that the assessment of 10 more MS patients per week (from a baseline of 16) was associated with lower risk of TI (adjusted coefficient −10.2; 95%CI −18.4 to −2.0).

## Aversion to Ambiguity and TI

For the primary outcome, high aversion to ambiguity in the financial domain was associated with TI (86.4 vs. 63.5%; $p = 0.042$). High ambiguity aversion in the health domain was not associated with TI (76.9 vs. 65.7%; adjusted OR 1.79, 95%CI 0.61–5.25). Multivariable logistic regression analysis showed that high aversion to ambiguity in the financial domain was the strongest predictor of TI, significantly stronger even than aversion to ambiguity in the health domain (adjusted OR 7.39; 1.40–38.9) (**Table 3**; Table S2 and Figures S2 and S3 in Supplementary Material). The results were also consistent when ambiguity aversion was defined by 50% unknown probability (adjusted OR 3.29; 1.21–8.99).

A sensitivity analysis showed that high aversion to ambiguity was also the strongest predictor of TI when applying the EMA criteria (adjusted OR 8.01; 95%CI 1.73–37.1) for the composite outcome of disease activity (modified Rio criteria) or evidence of progression (OR 4.41; 95%CI 1.04–18.7). The results remained consistent when models included all explanatory variables of interest, including number of patients seen per week (Table S2 in Supplementary Material).

Low tolerance to uncertainty (physician's reaction to uncertainty survey) was associated with higher prevalence of TI (85.4 vs. 56.4%; adjusted OR 4.73, 1.63–13.7) (**Figure 2**). The association between TI and low tolerance to uncertainty was independent of the association between TI and high ambiguity aversion (Table S2 in Supplementary Material).

**TABLE 2 | Prevalence of therapeutic inertia (TI) among multiple sclerosis (MS) specialists and general neurologists.**

| Outcome | MS specialists *N* = 64 (66.7) | General neurologists *N* = 32 (33.3) | *p*-Value |
|---|---|---|---|
| **TI (criterion)** | | | |
| Clinico-radiological | 40 (62.5) | 26 (81.3) | 0.062 |
| European Medicines Agency | 13 (20.3) | 15 (46.9) | 0.007 |
| Modified Rio or progression | 39 (60.9) | 26 (81.3) | 0.045 |

Conversely, willingness to take risk in multiple domains (as measured by the SOEP survey) or herding was not associated with TI (Table S2 in Supplementary Material).

# DISCUSSION

Multiple sclerosis patients and their treating physicians are routinely confronted with uncertainties concerning the diagnosis, prognosis, disease course, and disease-modifying therapies (27).

In the present study, we applied validated experiments and surveys from behavioral economics to evaluate the association between physicians' aversion to risk or ambiguity and TI (15). We found that TI affects nearly 7 out of 10 neurologists caring for MS patients but was less common among physicians with greater patient volumes per week or MS specialists. High aversion to ambiguity was the strongest predictor of TI even after adjusting for relevant confounders (e.g., age, practice setting, years in practice, percentage of time in clinical practice, overconfidence, time of survey completion, etc.). Lower tolerance to uncertainty was also associated with 3.5-fold higher risk of TI. On average, the assessment of 10 more MS patients per week was associated with lower risk of TI. Our results were consistent when employing various criteria on when to escalate therapy, as defined based on disease activity and/or progression. Physicians' characteristics (e.g., age, gender, practice setting, years in practice, percentage of time in clinical practice) were not associated with TI.

In the last few years, there has been an increase in the availability and variability of therapeutic options for the management of MS. Although having more options is perceived as beneficial, consumer studies in psychology suggest the higher the number of options, the more difficult the decision, leading to a higher number of less satisfactory choices (labeled as "choice overload") (34, 35).

Our results have practical implications that deserve comment. We showed that either a simple experiment or a short survey outside of the medical domain that assess aversion to ambiguity or tolerance to uncertainty may help to identify TI among neurologists and MS experts. The lack of escalation of therapies may lead to greater disability of MS patients, increasing the health-care costs and production losses due to incapacity to work. In Europe, the mean annual cost per person with MS has been estimated at €23,000 for EDSS score 0.0–3.5, rising as disability increases to €46,000 for EDSS score 4.0–6.5, and €77,000 for EDSS score 7.0–9.5 (36).

129

**TABLE 3 | Effect of high ambiguity aversion according to different definitions of therapeutic inertia (TI).**

| Outcome | Prevalence (%) of TI in the cohort | Adjusted model for ambiguity aversion[a] | | Adjusted model for ambiguity aversion[b] | |
|---|---|---|---|---|---|
| | | OR (95%CI) | *c*-Statistics | OR (95%CI) | *c*-Statistics |
| **TI (criterion)** | | | | | |
| Clinico-radiological | 66 (68.8) | 7.39 (1.40–38.9) | 0.804 | 8.01 (1.01–73.3) | 0.828 |
| European Medicines Agency | 28 (29.2) | 8.02 (1.37–37.1) | 0.777 | 7.17 (1.36–37.6) | 0.796 |
| Modified Rio or progression (Expanded Disability Status Scale >1) | 65 (67.7) | 4.41 (1.04–18.7) | 0.791 | 4.01 (0.83–19.3) | 0.811 |

[a]*Models derived from stepwise logistic regression with backward selection with p > 0.2 level for removal.*

[b]*Models derived from logistic regression including all variables of interest (age, sex, number of multiple sclerosis patients seen per week, practice setting, academic profile, risk aversion, ambiguity aversion, tolerance to uncertainty, herding, and overconfidence).*

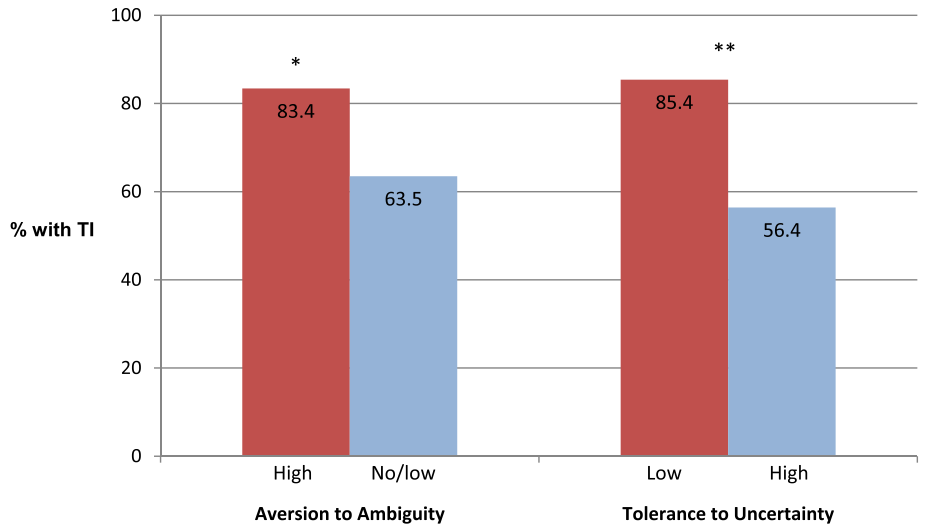



**FIGURE 2 | Prevalence of therapeutic inertia (TI) among participants with high ambiguity aversion in the financial domain and low tolerance to uncertainty in patient care**. See description in the text for the criteria of TI. *$p = 0.042$; **$p < 0.01$.

Factors associated with TI include low volume of MS patients, non-specialty, and physicians' ambiguity preferences (e.g., low tolerance to uncertainty in patient care, high aversion to financial ambiguity). Taken together, TI may be explained, at least in part, by (i) the aversion of neurologists to escalate treatment when the available options can have more serious side effects; (ii) the limited education (or experience) of neurologists regarding the risk profile of new DMTs, and (iii) participants' preference to continue with a known medication profile vs. the unknown risks of a new agent. Other studies have found that TI was associated with lack of training and clinical uncertainty (5). Physicians with better coping strategies and more tolerance to ambiguity may be more likely to choose optimal treatments leading to better patients' outcomes (37).

The prevailing significance of aversion to ambiguity in the financial over the health domain in the multivariate analysis may be related to either methodological differences when measuring each variable or suggest an underlying hardwired representation of aversion to ambiguity and TI that is easier elicited in the financial domain (18).

The results of DIScUTIR MS may not only be relevant for MS care but also be seen as the initial step to inform the design of a larger worldwide intervention, including physicians' high aversion to ambiguity in the financial domain and low tolerance to uncertainty in patient care when assessing the use of new agents.

Our study has limitations that deserve comment. First, the study was conducted in Spain exclusively, thus limiting the generalizability of our results to other cultural contexts. Moreover, some participants may have responded based on their current local restrictions to prescribe specific DMTs. However, cognitive distortions and risk preferences have been identified in several studies of physicians' decisions across the world and are thus probably not limited to a specific region or country (38). Second, the assessment of case scenarios may not fully capture decisions made in real clinical practice, even though specialists designed and recognized the scenarios as close to daily practice. In addition, participants may refer their MS patients to an MS outpatient clinic as part of a standard practice, which may have influenced our results. Third, considering the relatively low sample size, our findings should be viewed as exploratory. However, our results were consistent across several criteria of TI and adjusted models. Fourth, the concept and definition of TI applied to MS care is not widely disseminated and not yet generally accepted in MS

care. Nevertheless, we used a practical definition of TI (absence of escalation in the face of a clinical relapse plus evidence of imaging activity), which is supported by consensus panels, as well as by MS studies and other areas showing improvements in clinical outcomes when escalating therapies (i.e., blood pressure and diabetes) (14, 39, 40).

Despite these limitations, our study is the first step in the understanding of how specific characteristics of physicians (i.e., high ambiguity aversion, low tolerance to uncertainty) directly influence therapeutic decisions in MS patients beyond demographic factors, medical expertise, practice setting, patients' factors, or their treatment preferences. Using a novel approach that combines case-vignettes with the assessment of cognitive distortions through experiments from behavioral economics, we were able to expand our current understanding of decision-making under uncertainty in MS care.

Although MS experts have an expanded therapeutic arsenal compared to a decade ago, our study shows that nearly 7 out of 10 neurologists exhibited TI leading to suboptimal decisions. The results of DIScUTIR MS provide vital information to initiate discussions on behavioral strategies and incentives in order to ameliorate physicians' inertia to escalate therapies leading to better outcomes and quality of life for MS patients (41). For example, training in risk management and decision-making, as well as, educational interventions are needed to overcome knowledge-to-action gaps (and reduce the TI) in MS care. This is relevant considering the lack of well-established MS guidelines concerning clinical scenarios under uncertainty or progression of disease and the limited understanding on how physicians' preferences (e.g., aversion to ambiguity) have a global impact on medical and daily life decisions (42).

## ETHICS STATEMENT

All subjects gave consent (online) in accordance with the Declaration of Helsinki. The protocol was approved by the Research Ethics Board of St. Michael's Hospital, University of Toronto, Canada.

## AUTHOR CONTRIBUTIONS

GS: study concept and design, acquisition of data, analysis and interpretation of the data, and obtaining funding; AS: study concept and design, interpretation of the data, and critical revision of the manuscript for intellectual content; DP: interpretation of the data and critical revision of the manuscript for intellectual content; DS: study design, interpretation of the data, and critical revision of the manuscript for intellectual content; CR, JM, and PT: study concept and design, interpretation of the data, critical revision of the manuscript for intellectual content, and study supervision.

## ACKNOWLEDGMENTS

The authors are most grateful to all physicians participating in the study. The authors appreciate the support from the Department of Economics at the University of Zurich by facilitating access to Qualtrics platform.

## FUNDING

The study was sponsored by the Sociedad Española de Neurologia (SEN) and funded by an operating grant from Roche Farma, Spain. The authors acknowledge as well the grant CRSII3_141965 to PNT (Shared grant). The sponsors were not involved in the design, execution, analysis, and interpretation or reporting of the results.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://journal.frontiersin.org/article/10.3389/fneur.2017.00065/full#supplementary-material.

## REFERENCES

1. Glimcher P, Fehr E. *Neuroeconomics: Decision Making and the Brain*. 2nd ed. San Diego: Academic Press (2014). 2013 p.
2. Dijkstra IS, Pols J, Remmelts P, Brand PL. Preparedness for practice: a systematic cross-specialty evaluation of the alignment between postgraduate medical education and independent practice. *Med Teach* (2015) 37(2):153–61. doi:10.3109/0142159X.2014.929646
3. Platt ML, Huettel SA. Risky business: the neuroeconomics of decision making under uncertainty. *Nat Neurosci* (2008) 11(4):398–403. doi:10.1038/nn2062
4. Camerer C, Weber M. Recent developments in modeling preferences: uncertainty and ambiguity. *J Risk Uncertain* (1992) 5(4):325–70. doi:10.1007/BF00122575
5. Kerr EA, Zikmund-Fisher BJ, Klamerus ML, Subramanian U, Hogan MM, Hofer TP. The role of clinical uncertainty in treatment decisions for diabetic patients with uncontrolled blood pressure. *Ann Intern Med* (2008) 148(10):717–27. doi:10.7326/0003-4819-148-10-200805200-00004
6. Reach G. Clinical inertia, uncertainty and individualized guidelines. *Diabetes Metab* (2014) 40(4):241–5. doi:10.1016/j.diabet.2013.12.009
7. Noyes K, Weinstock-Guttman B. Impact of diagnosis and early treatment on the course of multiple sclerosis. *Am J Manag Care* (2013) 19(17 Suppl): s321–31.
8. Sormani MP, Rio J, Tintore M, Signori A, Li D, Cornelisse P, et al. Scoring treatment response in patients with relapsing multiple sclerosis. *Mult Scler* (2013) 19(5):605–12. doi:10.1177/1352458512460605
9. Duquette P, Giacomini PS, Bhan V, Hohol M, Schecter R. Balancing early aggression against risk of progression in multiple sclerosis. *Can J Neurol Sci* (2016) 43(1):33–43. doi:10.1017/cjn.2015.302
10. Prosperini L, Gianni C, Leonardi L, De Giglio L, Borriello G, Galgani S, et al. Escalation to natalizumab or switching among immunomodulators in relapsing multiple sclerosis. *Mult Scler* (2012) 18(1):64–71. doi:10.1177/1352458511417481
11. O'Connor PJ, Sperl-Hillen JAM, Johnson PE, Rush WA, Biltz G. Clinical inertia and outpatient medical errors. In: Henriksen K, Battles JB, Marks ES, Lewin DI, editors. *Advances in Patient Safety: From Research to Implementation (Volume 2: Concepts and Methodology)*. Rockville, MD: Agency for Healthcare Research and Quality (US) (2005). p. 293–308.
12. Mohan AV, Phillips LS. Clinical inertia and uncertainty in medicine. *JAMA* (2011) 306(4):383; author reply 383–4. doi:10.1001/jama.2011.1044
13. Phillips LS, Branch WT, Cook CB, Doyle JP, El-Kebbi IM, Gallina DL, et al. Clinical inertia. *Ann Intern Med* (2001) 135(9):825–34. doi:10.7326/0003-4819-135-9-200111060-00012
14. Okonofua EC, Simpson KN, Jesri A, Rehman SU, Durkalski VL, Egan BM. Therapeutic inertia is an impediment to achieving the Healthy People 2010 blood pressure control goals. *Hypertension* (2006) 47(3):345–51. doi:10.1161/01.HYP.0000200702.76436.4b

15. Saposnik G, Sempere AP, Raptis R, Prefasi D, Selchen D, Maurino J. Decision making under uncertainty, therapeutic inertia, and physicians' risk preferences in the management of multiple sclerosis (DIScUTIR MS). *BMC Neurol* (2016) 16(1):58. doi:10.1186/s12883-016-0577-4
16. Garcia Merino A, Ramon Ara Callizo J, Fernandez Fernandez O, Landete Pascual L, Moral Torres E, Rodriguez-Antiguedad Zarrantz A. Consensus statement on the treatment of multiple sclerosis by the Spanish Society of Neurology in 2016. *Neurologia* (2016) S0213-4853(16):30029–30029. doi:10.1016/j.nrl.2016.02.026
17. Saposnik G, Johnston SC. Decision making in acute stroke care: learning from neuroeconomics, neuromarketing, and poker players. *Stroke* (2014) 45(7):2144–50. doi:10.1161/STROKEAHA.114.005462
18. Levy I, Snell J, Nelson AJ, Rustichini A, Glimcher PW. Neural representation of subjective value under risk and ambiguity. *J Neurophysiol* (2010) 103(2):1036–47. doi:10.1152/jn.00853.2009
19. Anderson LR, Mellor JM. Predicting health behaviors with an experimental measure of risk preference. *J Health Econ* (2008) 27(5):1260–74. doi:10.1016/j.jhealeco.2008.05.011
20. Binmore K, Stewart L, Voorhoeve A. How much ambiguity aversion? Finding indifferences between Ellsberg's risky and ambiguous bets. *J Risk Uncertain* (2012) 45(3):215–38. doi:10.1007/s11166-012-9155-3
21. Gross CP, Vogel EW, Dhond AJ, Marple CB, Edwards RA, Hauch O, et al. Factors influencing physicians' reported use of anticoagulation therapy in nonvalvular atrial fibrillation: a cross-sectional survey. *Clin Ther* (2003) 25(6):1750–64. doi:10.1016/S0149-2918(03)80167-4
22. Christopoulos GI, Tobler PN, Bossaerts P, Dolan RJ, Schultz W. Neural correlates of value, risk, and risk aversion contributing to decision making under risk. *J Neurosci* (2009) 29(40):12574–83. doi:10.1523/JNEUROSCI.2614-09.2009
23. Dohmen T, Falk A, Huffman D, Sunde U, Schupp J, Wagner GG. Individual risk attitudes: measurement, determinants, and behavioral consequences. *J Eur Econ Assoc* (2011) 9(3):522–50. doi:10.1111/j.1542-4774.2011.01015.x
24. Gerrity MS, DeVellis RF, Earp JA. Physicians' reactions to uncertainty in patient care. A new measure and new insights. *Med Care* (1990) 28(8):724–36. doi:10.1097/00005650-199008000-00005
25. Gerrity M, White K, DeVellis R, Dittus R. Physicians' reactions to uncertainty: refining the constructs and scales. *Motiv Emot* (1995) 19(3):175–91. doi:10.1007/BF02250510
26. Prosperini L, Mancinelli CR, De Giglio L, De Angelis F, Barletta V, Pozzilli C. Interferon beta failure predicted by EMA criteria or isolated MRI activity in multiple sclerosis. *Mult Scler* (2014) 20(5):566–76. doi:10.1177/1352458513502399
27. Bermel RA, You X, Foulds P, Hyde R, Simon JH, Fisher E, et al. Predictors of long-term outcome in multiple sclerosis patients treated with interferon beta. *Ann Neurol* (2013) 73(1):95–103. doi:10.1002/ana.23758
28. Sormani MP, De Stefano N. Defining and scoring response to IFN-beta in multiple sclerosis. *Nat Rev Neurol* (2013) 9(9):504–12. doi:10.1038/nrneurol.2013.146
29. Sormani MP, Gasperini C, Romeo M, Rio J, Calabrese M, Cocco E, et al. Assessing response to interferon-beta in a multicenter dataset of patients with MS. *Neurology* (2016) 87(2):134–40. doi:10.1212/WNL.0000000000002830
30. Tramacere I, Del Giovane C, Salanti G, D'Amico R, Filippini G. Immunomodulators and immunosuppressants for relapsing-remitting multiple sclerosis: a network meta-analysis. *Cochrane Database Syst Rev* (2015) (9):CD011381. doi:10.1002/14651858.CD011381.pub2
31. Freedman MS, Selchen D, Arnold DL, Prat A, Banwell B, Yeung M, et al. Treatment optimization in MS: Canadian MS Working Group updated recommendations. *Can J Neurol Sci* (2013) 40(3):307–23. doi:10.1017/S0317167100014244
32. Wattjes MP, Rovira A, Miller D, Yousry TA, Sormani MP, de Stefano MP, et al. Evidence-based guidelines: MAGNIMS consensus guidelines on the use of MRI in multiple sclerosis – establishing disease prognosis and monitoring patients. *Nat Rev Neurol* (2015) 11(10):597–606. doi:10.1038/nrneurol.2015.157
33. Correale J, Abad P, Alvarenga R, Alves-Leon S, Armas E, Barahona J, et al. Management of relapsing-remitting multiple sclerosis in Latin America: practical recommendations for treatment optimization. *J Neurol Sci* (2014) 339(1–2):196–206. doi:10.1016/j.jns.2014.02.017
34. Iyengar SS, Lepper MR. When choice is demotivating: can one desire too much of a good thing? *J Pers Soc Psychol* (2000) 79(6):995–1006. doi:10.1037/0022-3514.79.6.995
35. Milte R, Ratcliffe J, Chen G, Lancsar E, Miller M, Crotty M. Cognitive overload? An exploration of the potential impact of cognitive functioning in discrete choice experiments with older people in health care. *Value Health* (2014) 17(5):655–9. doi:10.1016/j.jval.2014.05.005
36. Kobelt G, Berg J, Lindgren P, Fredrikson S, Jonsson B. Costs and quality of life of patients with multiple sclerosis in Europe. *J Neurol Neurosurg Psychiatry* (2006) 77(8):918–26. doi:10.1136/jnnp.2006.090365
37. Yee LM, Liu LY, Grobman WA. The relationship between obstetricians' cognitive and affective traits and their patients' delivery outcomes. *Am J Obstet Gynecol* (2014) 211(6):692.e1–6. doi:10.1016/j.ajog.2014.06.003
38. Blumenthal-Barby JS, Krieger H. Cognitive biases and heuristics in medical decision making: a critical review using a systematic search strategy. *Med Decis Making* (2015) 35(4):539–57. doi:10.1177/0272989X14547740
39. Huang LY, Shau WY, Yeh HL, Chen TT, Hsieh JY, Su S, et al. A model measuring therapeutic inertia and the associated factors among diabetes patients: a nationwide population-based study in Taiwan. *J Clin Pharmacol* (2015) 55(1):17–24. doi:10.1002/jcph.367
40. Goodin DS, Bates D. Treatment of early multiple sclerosis: the value of treatment initiation after a first clinical episode. *Mult Scler* (2009) 15(10):1175–82. doi:10.1177/1352458509107007
41. Emanuel EJ, Ubel PA, Kessler JB, Meyer G, Muller RW, Navathe AS, et al. Using behavioral economics to design physician incentives that deliver high-value care. *Ann Intern Med* (2016) 164(2):114–9. doi:10.7326/M15-1330
42. Lublin FD, Reingold SC, Cohen JA, Cutter GR, Sorensen PS, Thompson AJ, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology* (2014) 83(3):278–86. doi:10.1212/WNL.0000000000000560


**Conflict of Interest Statement:** PT and CR were funded by the Swiss National Science Foundation (PT: PP00P1_150739 and 00014_165884; CR: 105314_152891, CRSII3_141965, and 320030_143443). GS is supported by the Distinguished Clinicians Scientist Award from HSFC. AS, DP, DS, and JM have no disclosures.

**APPENDIX**

**Table S1. Criteria of Therapeutic inertia and case-scenarios**

| Criteria | Clinical + Radiological | EMA criteria | Composite modified Rio ≥2 or MS progression |
|---|---|---|---|
| **Case-scenario number** | | | |
| | 7 | 4 | 4 |
| | 8 | 8 | 7 |
| | 13 | 14 | 13 |
| | 14 | 15 | 14 |
| | | | 15 |
| **Definition** | Clinical relapse + at least 1 Gad enhancing lesion | 1 relapse last year + ≥ 9 new T2 OR ≥ 1 Gad T1 | MRI criterion: 1 if >5 new T2 lesions<br>Clinical criterion:<br>1 if 1 relapse<br>2 if ≥ 2 relapses over the first year or EDSS ≥1 from baseline |
| **Overlap with other measures** | **NO** | Herding | **NO** |

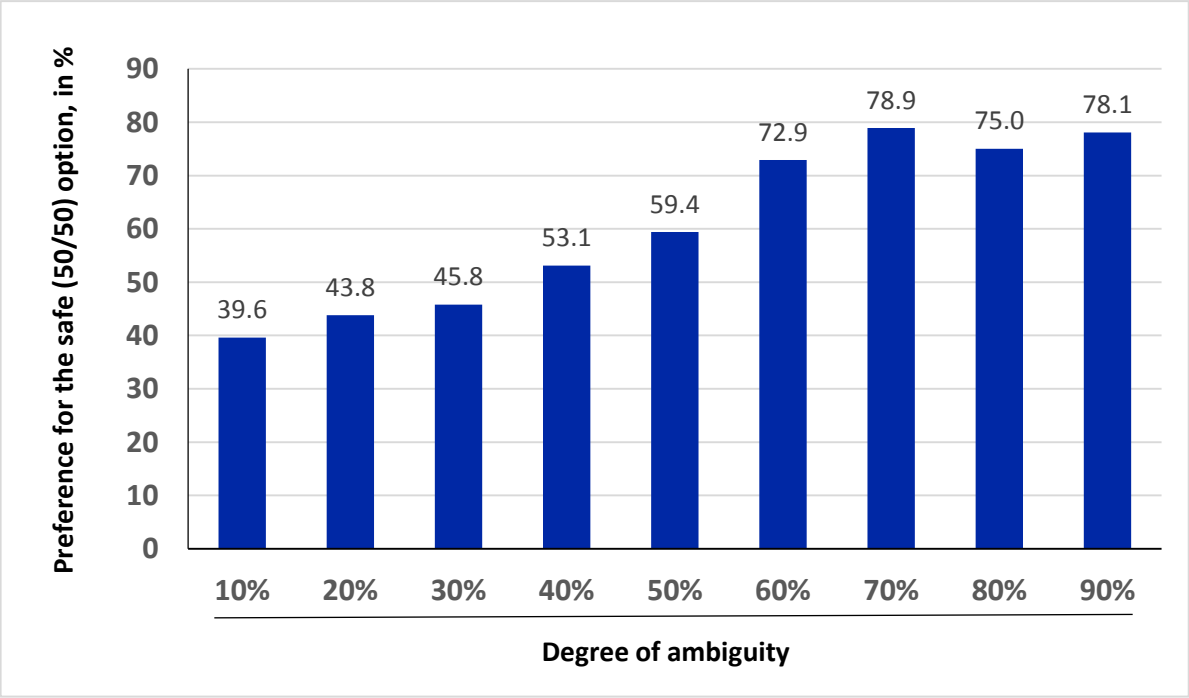
**Table S2. Variables associated with therapeutic inertia: full logistic regression models**

| | **Clinical + Radiological** | **EMA** | **Modified Rio or progression** |
|---|---|---|---|
| | OR; 95% CI | | |
| Ambiguity aversion-(Financial domain) | 8.6 (1.01-73.32) | 7.16; 1.36-37.63 | 3.91; 0.63-23.96 |
| Ambiguity aversion-(Health domain) | 0.75 (0.17-3.34) | 0.16; 0.30-0.92 | 1.10; 0.26-4.56 |
| Age | 0.93; 0.81-1.07 | 1.07; 0.84-1.36 | 0.95; 0.85-1.07 |
| Gender, male | 0.54; 0.18-1.67 | 0.85; 0.28-2.51 | 0.52; 0.17-1.56 |
| Years in practice | 1.02; 0.89-1.17 | 0.90; 0.69-1.17 | 1.01; 0.91-1.13 |
| Authorship | 1.18; 0.24-5.69 | 2.81; 0.73-10.74 | 1.37; 0.28-6.5 |
| Patients seen/week | 0.96; 0.93-1.00 | 0.96; 0.91-1.01 | 0.97; 0.93-1.00 |
| Setting, academic | 1.48; 0.41-5.28 | 0.44; 0.12-1.60 | 1.25; 0.36-4.31 |
| Attendance ECTRIMS 2015 | 1.38; 0.44-4.30 | 1.60; 0.54-4.77 | 1.26; 0.41-3.85 |
| SOEP | 0.74; 0.25-2.18 | 1.26; 0.43-3.70 | 0.87; 0.30-2.5 |
| Low tolerance to uncertainty | 3.9; 1.22-12.45 | 0.65; 0.21-1.93 | 4.27; 1.36-13.4 |
| Herding experiment | 0.37; 0.06-2.08 | 0.36; 0.09-1.37 | 0.34; 0.062-1.85 |
| Risk Aversion | 0.70; 0.18-2.68 | 2.35; 0.62-8.86 | 0.83; 0.22-3.11 |
| Overconfidence | 1.21; 0.30-4.89 | 1.20; 0.34-4.22 | 1.35; 0.34-5.37 |

**Figure S1. Representation of participants in DIScUTIR MS**

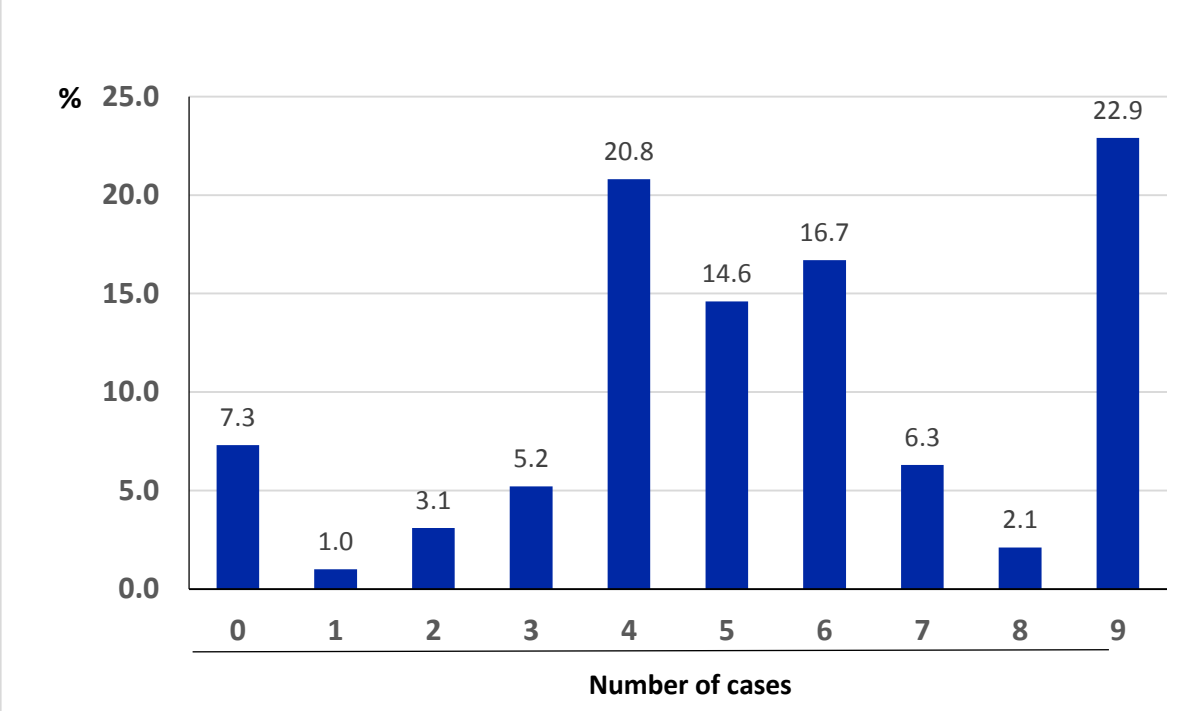

Participants from all territories in Spain were represented. Dot sizes represent the number of participants in each district.

**Figure S2. Preference for the known probability (50/50) option for each the nine scenarios with unknown probability (ambiguity aversion)**




Note the increasing number of participants choosing the known probability option (aversion to ambiguity) when the unknown probability increases.

**Figure S3. Number of times that participant choose the known probability (50/50) option over the ambiguous option**




Note: This figure represents the number of scenarios that participants choose the known probability option (50/50) over the ambiguous (unknown probability) option. For example, 20.8% of participants selected the 50/50 option in 4 scenarios, 16.7% selected the 50/50option in 6 scenarios. Overall, 22.9% of participants selected the known probability (50/50) option in all nine scenarios suggestive of complete aversion to ambiguity.

# C. Appendix to Study 3

JAMA Network | Open

Original Investigation | Neurology

## Comparison of Physician Therapeutic Inertia for Management of Patients With Multiple Sclerosis in Canada, Argentina, Chile, and Spain

Noora Almusalam, MD; Jiwon Oh, MD, PhD, FRCPC; Maria Terzaghi, RPh; Jorge Maurino, MD; Fabien Bakdache, PharmD; Alonso Montoya, MD; Fernando Caceres, MD; Gustavo Saposnik, MD, MSc, FRCPC

### Abstract

**IMPORTANCE** There is growing interest in understanding and addressing factors that govern the decision-making process in multiple sclerosis (MS) care. Therapeutic inertia (TI) is the failure to escalate therapy when goals are unmet. Limited data are available on the prevalence of TI and factors affecting therapeutic decisions in the management of patients with MS worldwide.

**OBJECTIVES** To compare TI across 4 countries (Canada, Argentina, Chile, and Spain) and to identify factors contributing to TI.

**DESIGN, SETTING, AND PARTICIPANTS** Prospective cohort study conducted between July 10, 2017, and May 4, 2018. Participants were exposed to behavioral experiments in which instruments were used to assess their risk preferences (eg, aversion to ambiguity) and therapeutic decisions in 10 simulated MS case scenarios. Mixed-effects linear and logistic regression analyses were performed to determine the association between the participants' baseline characteristics and TI. The association of unmeasured confounders was assessed by the E-value and a bootstrapping analysis. This multicenter study included neurologists practicing at academic and community centers in Canada, Argentina, Chile, and Spain who make therapeutic decisions for patients with MS.

**MAIN OUTCOMES AND MEASURES** The primary outcome was the prevalence of TI. The TI score was calculated by dividing the number of case scenarios in which participants showed TI by the number of case scenarios that measured TI. Higher TI scores indicated greater degrees of TI. The secondary outcome was the identification of factors that contributed to TI.

### Key Points

**Question** What is the prevalence of therapeutic inertia and its associated factors?

**Findings** In this cohort study that included 195 neurologists with expertise in multiple sclerosis, the prevalence and magnitude of therapeutic inertia among a Canadian group were the lowest compared with Argentina, Chile, and Spain. Seeing a higher number of patients per week, years of practice, and participation from Canada were associated with a lower likelihood of therapeutic inertia, whereas aversion to ambiguity was associated with a higher likelihood of therapeutic inertia.

**Meaning** Therapeutic inertia is common among practicing neurologists, with practical implications for patients with multiple sclerosis.

This paper is published as Almusalam N, Oh J, Terzaghi M, Maurino J, Bakdache F, Montoya A, Caceres F, Saposnik G. Comparison of Physician Therapeutic Inertia for Management of Patients With Multiple Sclerosis in Canada, Argentina, Chile, and Spain. JAMA Netw Open.;2(7):e197093, 2019.

JAMA Network Open

Original Investigation | Neurology

# Comparison of Physician Therapeutic Inertia for Management of Patients With Multiple Sclerosis in Canada, Argentina, Chile, and Spain

Noora Almusalam, MD; Jiwon Oh, MD, PhD, FRCPC; Maria Terzaghi, RPh; Jorge Maurino, MD; Fabien Bakdache, PharmD; Alonso Montoya, MD; Fernando Caceres, MD; Gustavo Saposnik, MD, MSc, FRCPC

## Abstract

**IMPORTANCE** There is growing interest in understanding and addressing factors that govern the decision-making process in multiple sclerosis (MS) care. Therapeutic inertia (TI) is the failure to escalate therapy when goals are unmet. Limited data are available on the prevalence of TI and factors affecting therapeutic decisions in the management of patients with MS worldwide.

**OBJECTIVES** To compare TI across 4 countries (Canada, Argentina, Chile, and Spain) and to identify factors contributing to TI.

**DESIGN, SETTING, AND PARTICIPANTS** Prospective cohort study conducted between July 10, 2017, and May 4, 2018. Participants were exposed to behavioral experiments in which instruments were used to assess their risk preferences (eg, aversion to ambiguity) and therapeutic decisions in 10 simulated MS case scenarios. Mixed-effects linear and logistic regression analyses were performed to determine the association between the participants' baseline characteristics and TI. The association of unmeasured confounders was assessed by the E-value and a bootstrapping analysis. This multicenter study included neurologists practicing at academic and community centers in Canada, Argentina, Chile, and Spain who make therapeutic decisions for patients with MS.

**MAIN OUTCOMES AND MEASURES** The primary outcome was the prevalence of TI. The TI score was calculated by dividing the number of case scenarios in which participants showed TI by the number of case scenarios that measured TI. Higher TI scores indicated greater degrees of TI. The secondary outcome was the identification of factors that contributed to TI.

**RESULTS** Of 300 neurologists with expertise in MS care who were invited to be part of the study, 226 (75.3%) agreed to participate. Among those who initially showed interest in participating, 195 physicians (86.3%) completed the study, while 31 did not. The mean (SD) age of participants was 43.3 (11.2) years; 52.3% were male. Therapeutic inertia was present in 72.8% (142 of 195) of participants, leading to suboptimal decisions in 20.4% (318 of 1560) of case scenarios. The prevalence of TI among the Canadian group was the lowest compared with the other 3 countries (60.0% [33 of 55] vs 77.9% [109 of 140]; $P = .01$). For the primary outcome, the TI score in the Canadian group (mean [SD], 0.98 [1.15]) was significantly lower compared with groups from other countries (mean [SD], 1.70 [1.43] for Argentina, 2.24 [1.54] for Chile, and 2.56 [1.64] for Spain) ($P = .001$). The mixed-effects linear models revealed that participants from Argentina, Chile, and Spain (combined) had higher TI scores compared with their Canadian counterparts (β coefficient, 0.90; 95% CI, 0.52-1.28; $P < .001$). A higher number of patients with MS per week (OR, 0.44; 95% CI, 0.22-0.88), years of practice (OR, 0.93; 95% CI, 0.86-0.99), and participation from Canada (OR, 0.47; 95% CI, 0.23-0.96) were associated with a lower likelihood of TI. Aversion to ambiguity was

(continued)

## Key Points

**Question** What is the prevalence of therapeutic inertia and its associated factors?

**Findings** In this cohort study that included 195 neurologists with expertise in multiple sclerosis, the prevalence and magnitude of therapeutic inertia among a Canadian group were the lowest compared with Argentina, Chile, and Spain. Seeing a higher number of patients per week, years of practice, and participation from Canada were associated with a lower likelihood of therapeutic inertia, whereas aversion to ambiguity was associated with a higher likelihood of therapeutic inertia.

**Meaning** Therapeutic inertia is common among practicing neurologists, with practical implications for patients with multiple sclerosis.

+ Supplemental content

Author affiliations and article information are listed at the end of this article.

*Abstract (continued)*

associated with a 2-fold higher likelihood of TI (OR, 2.25; 95% CI, 1.02-5.00). All 95% CIs of the β coefficients of covariates were lower than the E-value of 2.35, making it unlikely for the results to be due to the association of unmeasured confounders.

**CONCLUSIONS AND RELEVANCE** This study showed that Canadian participants had the lowest prevalence and magnitude of TI. Higher TI scores were associated with a lower expertise in MS care and with a greater tendency for aversion to ambiguity.

*JAMA Network Open.* 2019;2(7):e197093. doi:10.1001/jamanetworkopen.2019.7093

## Introduction

Multiple sclerosis (MS) is an evolving field, with an increasing number of proven effective therapies.[1-3] Given the broad spectrum of disease-modifying therapies (DMTs), neurologists may face therapeutic dilemmas when individualizing decisions regarding the most appropriate DMT for their patients.[4,5] It is challenging to decide on an optimal DMT for an individual patient based on the patient's clinical disease activity, magnetic resonance imaging (MRI) lesion burden, drug availability, adverse effect profile, and patient preferences.[1] Despite the recent advances in MS therapeutics, many patients remain undertreated.[6] Numerous factors contribute to the suboptimal management, including education gaps in both risk management and decision making among treating physicians.[7]

The concept of clinical inertia was initially introduced by Phillips et al[8] in 2001, who defined it as lack of therapy escalation when it is clinically indicated. The term *clinical inertia* was substituted with *therapeutic inertia* (TI) in 2006 by Okonofua et al.[9] Therapeutic inertia is prevalent not only among patients with MS but also in other chronic conditions, such as diabetes and hypertension.[10-12]

There is growing interest in understanding and addressing factors that govern the decision-making process[13,14] in MS. A particular interest is identifying variables that alter the therapeutic decision making of remaining on the same DMT or escalating to a more effective agent, which can be associated with a possible increase in the risk of serious adverse effects.[1,2]

Factors contributing to TI remain poorly understood. However, the results of some studies[6,15] suggest that physician-associated factors (eg, low tolerance to uncertainty and aversion to ambiguity) may have a role in the decision-making process. In addition, there is limited understanding of country-specific differences with regard to the prevalence of TI. Accordingly, the primary aim of this study was to compare the prevalence of TI among MS-treating physicians in 4 countries (Canada, Argentina, Chile, and Spain). The secondary aim of the study was to identify factors contributing to TI.

## Methods

### Study Population

This prospective cohort study was conducted between July 10, 2017, and May 4, 2018, and followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline. Our study population consisted of neurologists who primarily focus their clinical practice on MS care from 4 countries (55 in Canada, 90 in Argentina, 25 in Chile, and 25 in Spain). Participants were invited to take part in the study via email by scientific organizations based in each country (Canadian Consortium of MS Clinics and NeuroSens, Instituto de Neurociencias Buenos Aires [INEBA] and Argentine Neurological Society, Chilean Society of Neurology, and Spanish Neurological Society). Each participant received an email with instructions and a link to start the study. Participants not actively involved in patient care or those who followed up a low volume of patients with MS (<12 patients with MS per year) were excluded.

Informed consent was provided online at the beginning of the study by each participant. Participants received compensation for completing the surveys that is equivalent to US $100. The study was approved by the Research Ethics Board of St Michael's Hospital, University of Toronto, Toronto, Ontario, Canada.

## Data Collection

Participants provided information regarding their clinical practice and expertise and completed behavioral experiments to assess their risk preferences. Thereafter, participants were exposed to 10 simulated MS case scenarios and made therapeutic choices (case scenarios are available in the eAppendix in the Supplement). Case scenarios were designed by team members (M.T. and G.S.) and by MS experts (J.O. and a nonauthor) to reflect current clinical practice. Eight of the cases required escalation of therapy; therefore, the failure to do so by the participant was considered as TI. The other 2 cases were designed to detect overtreatment when treatment escalation was not deemed medically necessary. All participants were exposed to the same case scenarios. Details of behavioral experiments in which instruments were used in the assessment of tolerance to uncertainty and risk aversion are described elsewhere.[6,16,17] In brief, participants were asked to choose between 2 options of (1) either winning US $400 or $0 when the probability is 50/50 (represented by a blue/red bar) vs (2) an option of unknown probability (represented by a blue/red bar covered by a gray bar) of the same outcome. Participants who favored the known probability of 50/50 were deemed to have aversion to ambiguity, while other participants were considered as having tolerance to uncertainty. Risk was assessed by asking participants to provide the minimal amount of US dollars (or healthy years) that they would prefer over a 50/50 chance of winning US $400 (or longer survival with adverse effects of a treatment). The degree of risk aversion of each individual corresponded to the difference of the expected value of the risky option (US $200) minus the participant's response.[16]

We also evaluated physicians' tolerance to uncertainty in a patient's care using the Reaction to Uncertainty Test.[17] The test was composed of 5 questions that the respondent rated from 0 to 5, which were summed to give a total score. Higher scores represent a lower tolerance to uncertainty. Low tolerance to uncertainty was defined as values above the median of the total score. Further details of the protocol are published elsewhere.[18] Responses from case scenarios were analyzed in light of responses from the behavioral components.

## Definitions

Disease activity was defined as a clinical relapse that was associated with the presence of one of the following MRI findings: at least 1 gadolinium-enhancing lesion or 5 or more new brain lesions.[19,20] These thresholds were defined based on prior studies demonstrating in patients receiving interferon beta that risk of treatment failure highly correlates with a clinical relapse and MRI lesions as defined above.[21] Disease progression in MS was defined as at least 1 point of sustained worsening from baseline in the Expanded Disability Status Scale Score.[22]

At the time of the study, treatment options for relapsing-remitting MS included first-line therapies (beta interferons, glatiramer acetate, teriflunomide, and dimethyl fumarate), second-line therapy (fingolimod), and third-line therapies (natalizumab and alemtuzumab).[4] Given variation in market approval status of some DMTs among the countries in which the study was conducted, ocrelizumab and cladribine were not included in case scenarios. For this analysis, we used the 3-line treatment escalation scheme according to current clinical practice.[6,23]

## Outcome Measures

The primary outcome was the prevalence of TI. As in previous studies,[6,24] the TI score was calculated by dividing the number of case scenarios in which participants showed TI by the number of case scenarios that measured TI (n = 8). Higher TI scores indicated greater degrees of TI. The TI score ranged between 0 and 8.

Therapeutic inertia prevalence was defined as the proportion of participants with TI in at least 1 case scenario. The secondary outcome was the identification of factors that contributed to TI.

### Statistical Analysis

The primary analysis compared TI between Canada and the other 3 countries combined (ie, Argentina, Chile, and Spain). We included different measures of TI as both continuous and categorical variables to determine the consistency of our results. We used mixed-effects linear and logistic models to assess associations between TI score and TI to determine the association between the participants' baseline characteristics and the primary outcome of interest after accounting for clustering. We included the following explanatory variables: age, sex, specialty, number of patients with MS seen per week, proportion of time devoted to clinical care, coauthor of a peer-reviewed publication in the last year (yes or no), practice setting (academic vs nonacademic), risk aversion, aversion to ambiguity, and physicians' reaction to uncertainty (above or below the median).

This multicenter study included neurologists practicing at academic and community centers in Canada, Argentina, Chile, and Spain who make therapeutic decisions for patients with MS. To account for unmeasured confounding, we used an E-value estimate and a bootstrapping analysis to compare the β coefficients in the normal, percentile, bias-corrected, and bias-corrected and accelerated 95% CIs. The E-value is a tool to assess the consequences of unmeasured confounding in observational studies.[25] By comparing the upper 95% CI with the 95% CI of covariates included in the models, the E-value provides an estimate of the residual confounding that could explain an observed association if an unmeasured covariate exists.[25]

All tests were 2 tailed, and *P* values less than .05 were considered statistically significant. The area under the curve was used to assess the discrimination, and the Hosmer-Lemeshow test was used to assess the calibration of the model.

## Results

Of 300 neurologists with expertise in MS care who were invited to be part of the study, 226 (75.3%) agreed to participate. Among those who initially showed interest in participating, 195 physicians (86.3%) completed the study, while 31 did not. The mean (SD) age of participants was 43.3 (11.2) years; 52.3% were male. Eighty-six (44.1%) were MS specialists. **Table 1** summarizes baseline characteristics by country.

Overall, the prevalence of TI was 72.8% (142 of 195), leading to suboptimal decisions in 20.4% (318 of 1560) of case scenarios. The mean (SD) TI score for the accountable 8 case scenarios was 1.68 (1.50). For every 10 case scenarios with moderate to high risk of disease progression, this suggested

Table 1. Baseline Characteristics of Participants Between July 10, 2017, and May 4, 2018

| Characteristic | Total (N = 195) | Canada (n = 55) | Combined (n = 140)[a] | *P* Value |
|---|---|---|---|---|
| Age, mean (SD), y | 43.3 (11.2) | 41.8 (12.0) | 43.9 (10.8) | .24 |
| Sex, No. (%) | | | | |
| Female | 93 (47.7) | 26 (47.3) | 67 (47.9) | .94 |
| Male | 102 (52.3) | 29 (52.7) | 73 (52.1) | |
| Specialty, No. (%) | | | | |
| General neurologist who cares for MS | 109 (55.9) | 24 (43.6) | 85 (60.7) | .03 |
| MS specialist | 86 (44.1) | 31 (56.4) | 55 (39.3) | |
| No. of patients with MS seen/week, mean (SD) | 19.4 (11.6) | 22.2 (14.6) | 18.4 (10.0) | .04 |
| Years of practice, mean (SD) | 16.7 (11.4) | 13.3 (11.2) | 17.9 (12.2) | .01 |
| ≥75% of time devoted to clinical practice, No. (%) | 98 (50.3) | 30 (54.5) | 68 (48.6) | .45 |
| Coauthor of a peer-reviewed publication in the last year, No. (%) | 92 (47.2) | 31 (56.4) | 61 (43.6) | .11 |

Abbreviation: MS, multiple sclerosis.

[a] Combined countries are Argentina, Chile, and Spain.

that there would be 2 suboptimal decisions (eg, lack of treatment escalation) when clinical evidence of relapses and radiological evidence of disease activity exist.

## Outcome Measures

### Comparison of TI Between Canada and Other Countries

For the primary outcome, the TI score in the Canadian group (mean [SD], 0.98 [1.15]) was significantly lower compared with groups from other countries (mean [SD], 1.70 [1.43] for Argentina, 2.24 [1.54] for Chile, and 2.56 [1.64] for Spain) ($P$ = .001). The prevalence of TI among the Canadian group was also the lowest compared with the other 3 countries (60.0% [33 of 55] vs 77.9% [109 of 140], $P$ = .01). Values adjusted for age, specialty, number of patients with MS seen per week, years of practice, and aversion to ambiguity are summarized in **Table 2** and **Figure 1**.

Participant's risk preferences based on assessments with behavioral instruments are listed in **Table 3**. Compared with Canada, participants from the other countries were more risk averse and had a lower tolerance to uncertainty (reflected by higher scores) (Table 3). For example, the Canadian group had a lower proportion of participants with values below the certainty equivalence of US $120

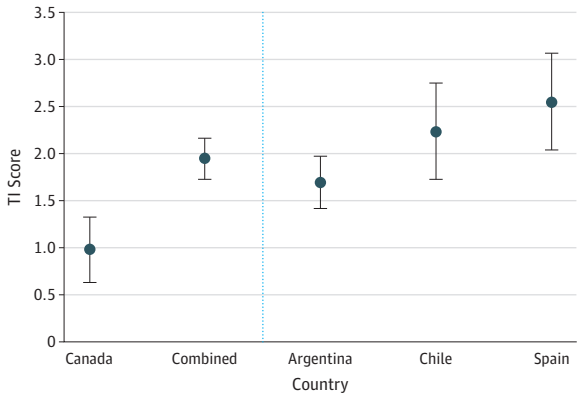
Table 2. Adjusted TI Score and Prevalence of TI by Country

| Country | TI Score (95% CI)[a] | Multivariable Analysis for TI Score, β Coefficient (95% CI) | *P* Value[a] | Prevalence of TI, % (95% CI)[a] |
|---|---|---|---|---|
| Canada | 0.98 (0.63-1.33) | 1 [Reference] | NA | 63.7 (49.4-76.0) |
| Argentina, Chile, and Spain combined | 1.95 (1.73-2.17) | 0.90 (0.52-1.28) | <.001 | 78.0 (70.0-84.2) |
| Argentina | 1.70 (1.42-1.98) | 0.65 (0.24-1.06) | .002 | 72.5 (62.9-82.2) |
| Chile | 2.24 (1.72-2.76) | 0.92 (0.21-1.64) | .01 | 61.7 (39.3-84.1) |
| Spain | 2.56 (2.04-3.08) | 1.74 (1.07-2.42) | <.001 | 96.4 (89.5-100) |

Abbreviations: NA, not applicable; TI, therapeutic inertia.

[a] Adjusted for age, specialty, number of patients with multiple sclerosis seen per week, years of practice, and aversion to ambiguity after accounting for clustering.

Figure 1. Therapeutic Inertia (TI) Score by Country




Adjusted TI scores are compared across the studied countries. The vertical bars represent 95% CIs. Estimates were derived from linear regression models after adjustment for age, specialty, number of patients with multiple sclerosis seen per week, years of practice, and aversion to ambiguity.

Table 3. Results of Behavioral Experiments Designed to Assess Risk and Aversion to Ambiguity[a]

| Country | Risk Financial Domain, Mean (SD) Score | Risk Health Domain, Mean (SD) Score | Aversion to Ambiguity, No./Total No. (%) | Physicians' Reaction to Uncertainty, Mean (SD) Score[b] |
|---|---|---|---|---|
| Canada | 213.5 (77.2) | 14.9 (3.4) | 41/55 (74.5) | 19.1 (8.6) |
| Argentina, Chile, and Spain combined | 179.5 (97.0)[c] | 14.3 (4.9) | 93/140 (66.4) | 23.1 (9.3)[d] |
| Argentina | 159.6 (102.3)[e] | 12.7 (7.0)[c] | 54/90 (60.0) | 23.3 (9.3)[e] |
| Chile | 233.2 (104.0) | 11.2 (5.1)[e] | 20/25 (80.0) | 23.1 (10.6) |
| Spain | 183.2 (47.9) | 14.9 (3.7) | 19/25 (76.0) | 22.5 (7.9) |

[a] *P* values are for comparison with Canada (reference group).

[b] Higher scores indicate a lower tolerance to uncertainty.

[c] $P$ = .02.

[d] $P$ = .006.

[e] $P$ < .001.

(value that identifies participants with risk aversion) compared with their counterparts (20.0% [11 of 55] vs 37.1% [52 of 140], *P* = .02).

Participants with an aversion to ambiguity had higher TI scores, although this factor did not reach statistical significance (mean, 1.90 vs 1.57; *P* = .15). Participants with an aversion to ambiguity also had a significantly higher TI prevalence (82.0% [50 of 61] vs 68.7% [92 of 134], *P* = .047).

The mixed-effects linear models revealed that participants from Argentina, Chile, and Spain (combined) had higher TI scores compared with their Canadian counterparts (β coefficient, 0.90; 95% CI, 0.52-1.28; *P* < .001). Similarly, participants from Argentina, Chile, and Spain (combined) had a higher likelihood of TI compared with the Canadian participants (odds ratio [OR], 2.30; 95% CI, 1.10-4.82; *P* = .03). The observed vs predicted TI scores after adjustment for covariates are shown in the eFigure in the Supplement. Details by country are listed in Table 2.

### Factors Associated With TI

The multivariable analysis aimed at identifying factors associated with the prevalence of TI revealed that a higher number of patients with MS per week (OR, 0.44; 95% CI, 0.22-0.88), years of practice (OR, 0.93; 95% CI, 0.86-0.99), and participation from Canada (OR, 0.47; 95% CI, 0.23-0.96) were associated with a lower likelihood of TI. Aversion to ambiguity was associated with a 2-fold higher likelihood of TI (OR, 2.25; 95% CI, 1.02-5.00) (**Figure 2**). The model showed good discrimination (area under the curve, 0.783) and calibration (Hosmer-Lemeshow test *P* = .90). Participants from Argentina, Chile, and Spain (combined) had 2.40 (95% CI, 1.16-5.06) higher odds of TI.

## Assessment of Unmeasured Confounding

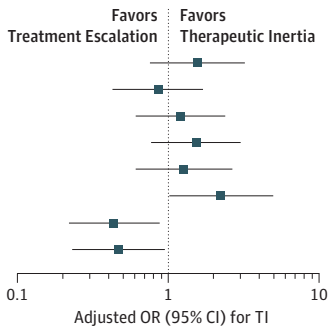
We estimated the E-value using the standardized difference of the mean TI score (point estimate, 0.667; standard error, 0.128; and true association with shift estimate, 0.227). The E-value for the point estimate was 2.35, and the CI was 1.66. All 95% CIs of the β coefficients of covariates were lower than the E-value of 2.35 (eTable 1 in the Supplement), making it unlikely for the results to be due to the association of unmeasured confounders. The bootstrapping analysis showed similarities among the β coefficients within each covariate, suggesting unbiased estimates of the TI scores (eTable 2 in the Supplement).

## Discussion

Therapeutic inertia is a well-known phenomenon influencing physicians who manage patients with chronic conditions, including MS.[6] Our study showed that neurologists with expertise in MS from 4 different countries exhibited varying degrees of TI. Overall, TI was present in 72.8% (142 of 195) of participants and altered at least 1 in 5 therapeutic decisions. Therapeutic inertia was 2.3-fold more prevalent among neurologists from Argentina, Chile, and Spain (combined) compared with their Canadian counterparts.

Figure 2. Factors Associated With Therapeutic Inertia (TI)

| Factors Associated With Therapeutic Inertia | OR (95% CI) |
|---|---|
| Age ≥45 y vs ≤44 y | 1.58 (0.76-3.26) |
| Sex male vs female | 0.86 (0.43-1.71) |
| Time devoted to clinical care ≥75% vs ≤74% | 1.21 (0.61-2.38) |
| Tolerance to uncertainty ≥23 vs ≤22 | 1.54 (0.78-3.05) |
| Risk financial domain ≥200 vs ≤199 | 1.28 (0.61-2.68) |
| Aversion to ambiguity present vs absent | 2.25 (1.02-5.00) |
| No. of patients with MS seen per week ≥22 vs ≤21 | 0.44 (0.22-0.88) |
| Canada vs other countries | 0.47 (0.23-0.96) |




Factors associated with TI are shown. A higher number of patients with multiple sclerosis (MS) seen per week and country (ie, Canada) were associated with a lower TI, whereas aversion to ambiguity was associated with a higher TI. OR indicates odds ratio.

The most important factors associated with TI include a lower expertise in MS care (eg, fewer years of experience and a lower number of patients with MS seen per week), country of practice, and a higher aversion to ambiguity. The mixed multivariable analysis supports the contention that aversion to ambiguity was an independent predictor of the prevalence of TI after accounting for demographic factors, level of expertise, and regional variations. We also found that unmeasured confounders were unlikely to have altered our findings.

The country of primary clinical practice was also identified as an independent predictor of TI. The observed differences among countries can be partially explained by variations in adherence to published MS management guidelines,[2,26,27] regional differences regarding eligibility and access to DMTs,[28] public funding of DMTs, and physician-related factors, such as risk preferences and education in risk management and decision making.

Recognizing the presence of TI and contributing factors is essential in identifying strategies aimed at improving medical education, which could lead to better patient outcomes. For example, the results of recent studies suggest that innovative therapeutic interventions (eg, a traffic light system) may be useful to ameliorate the prevalence and magnitude of TI. In brief, the traffic light system creates a link between a color (representing a risk level) and an action. Red light indicates "high risk" or "stop and think," yellow light indicates "intermediate risk" or "reassess soon," and green light indicates "low risk" or "continue the same strategy." When applied in clinical sittings, this method facilitates the decision-making process.[29] A subsequent randomized clinical trial targeting neurologists with MS expertise showed a 70% reduction in TI (OR, 0.30; 95% CI, 0.10-0.89) for the traffic light system educational intervention arm compared with usual care.[30]

### Limitations

Our study has several limitations. The small numbers of participants from Chile and Spain may have altered the accuracy of estimates. As a result, we reported the comparison of TI between Canada and the other 3 countries combined and also analyzed the role of unmeasured confounders. Another limitation is that physicians' decisions to escalate therapy are influenced by factors like availability of the drug, local policies, drug costs, variations in treatment guidelines in different countries, and other unmeasured variables that can alter the assessment of TI.[26,31,32] However, all included countries shared private and government-funded MS drug coverage in the absence of private health insurance.[33,34] A previous study[35] showed a limited role of costs in explaining therapeutic decisions and TI. In addition, physicians' performance on case scenarios might not accurately mirror real-life decisions, although such inconsistency would be expected to underestimate the true prevalence of TI among participants. Also, we cannot rule out the possibility of some residual confounding despite a comprehensive adjustment of baseline characteristics.

## Conclusions

This study demonstrates the high prevalence of and factors associated with TI even among neurologists with expertise in MS care. To date, our study is the first, to our knowledge, to systematically compare TI using identical case scenarios across countries. This study constitutes a first step in understanding the mechanisms of TI and increases awareness of its high prevalence. Our findings may also lead to the development of further studies that assess strategies to reduce TI, which may result in improved outcomes for patients with MS. We propose larger studies to evaluate the potential benefits of educational interventions to ameliorate TI. If proven effective, these strategies could be included in curricula for undergraduate and postgraduate medical programs to improve the existing education gaps of formal training in risk management and decision-making.

## ARTICLE INFORMATION

**Accepted for Publication:** May 23, 2019.

**Published:** July 17, 2019. doi:10.1001/jamanetworkopen.2019.7093



**Corresponding Author:** Gustavo Saposnik, MD, MSc, FRCPC, Division of Neurology, Department of Medicine, St Michael's Hospital, University of Toronto, 55 Queen St E, Toronto, ON M5C 1R6, Canada (saposnikg@smh.ca).

**Author Affiliations:** Division of Neurology, Department of Medicine, St Michael's Hospital, University of Toronto, Toronto, Ontario, Canada (Almusalam, Oh, Saposnik); Decision Neuroscience Unit, Li Ka Shing Knowledge Institute, St Michael's Hospital, University of Toronto, Toronto, Ontario, Canada (Terzaghi, Saposnik); Neuroscience Area, Medical Department, Roche Farma, Madrid, Spain (Maurino); Neuroscience, Medical Affairs, F. Hoffmann-La Roche Ltd, Mississauga, Ontario, Canada (Bakdache, Montoya); Instituto de Neurociencias Buenos Aires (INEBA), Buenos Aires, Argentina (Caceres); Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Zurich, Switzerland (Saposnik).

**Author Contributions:** Drs Almusalam and Saposnik had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

*Concept and design:* Almusalam, Maurino, Saposnik.

*Acquisition, analysis, or interpretation of data:* Oh, Terzaghi, Maurino, Bakdache, Montoya, Caceres, Saposnik.

*Drafting of the manuscript:* Oh, Terzaghi, Saposnik.

*Critical revision of the manuscript for important intellectual content:* Almusalam, Oh, Maurino, Bakdache, Montoya, Caceres, Saposnik.

*Statistical analysis:* Saposnik.

*Obtained funding:* Bakdache, Montoya, Saposnik.

*Administrative, technical, or material support:* Almusalam, Terzaghi, Bakdache, Caceres, Saposnik.

*Supervision:* Oh, Maurino, Montoya, Saposnik.

**Conflict of Interest Disclosures:** Dr Oh reported receiving research funding from the MS Society of Canada, the National Multiple Sclerosis Society, Brain Canada, Sanofi Genzyme, and Biogen; reported receiving personal compensation for consulting or speaking from EMD Serono, Sanofi Genzyme, Biogen, Novartis, Celgene, and Roche; and reported receiving grants and/or personal fees from Roche, Biogen Idec, Sanofi Genzyme, Celgene, and EMD Serono. Dr Maurino reported receiving support from Roche Farma. Dr Bakdache reported receiving personal fees from F. Hoffmann-La Roche Ltd. Dr Montoya reported receiving personal fees from Roche Canada. Dr Caceres reported receiving grants and/or personal fees from Roche, Biogen, Novartis, Merck, and Sanofi Genzyme. Dr Saposnik reported receiving grants and personal fees from Roche and reported being supported by the Heart and Stroke Foundation of Canada Career Award. No other disclosures were reported.

**Funding/Support:** The study was sponsored by the Instituto de Neurociencias Buenos Aires (INEBA), the Spanish Society of Neurology (SEN) with endorsement by the Sociedad Argentina de Neurologia (SNA), the Division of Neurology at the University of Toronto, and NeuroEconSolutions and was funded by an operating grant from Roche.

**Role of the Funder/Sponsor:** The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

**Additional Contributions:** We thank the physicians who participated in the study. We acknowledge support from the Department of Economics at the University of Zurich.

## REFERENCES

1. Montalban X, Gold R, Thompson AJ, et al. ECTRIMS/EAN guideline on the pharmacological treatment of people with multiple sclerosis. *Mult Scler*. 2018;24(2):96-120. doi:10.1177/1352458517751049

2. Rae-Grant A, Day GS, Marrie RA, et al. Practice guideline recommendations summary: disease-modifying therapies for adults with multiple sclerosis: report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. *Neurology*. 2018;90(17):777-788. doi:10.1212/WNL.0000000000005347

3. Giovannoni G. Disease-modifying treatments for early and advanced multiple sclerosis: a new treatment paradigm. *Curr Opin Neurol*. 2018;31(3):233-243. doi:10.1097/WCO.0000000000000561

**4**. Tramacere I, Del Giovane C, Salanti G, D'Amico R, Filippini G. Immunomodulators and immunosuppressants for relapsing-remitting multiple sclerosis: a network meta-analysis. *Cochrane Database Syst Rev*. 2015;(9):CD011381. doi:10.1002/14651858.CD011381.pub2

**5**. Alping P, Frisell T, Novakova L, et al. Rituximab versus fingolimod after natalizumab in multiple sclerosis patients. *Ann Neurol*. 2016;79(6):950-958. doi:10.1002/ana.24651

**6**. Saposnik G, Sempere AP, Prefasi D, et al. Decision-making in multiple sclerosis: the role of aversion to ambiguity for therapeutic inertia among neurologists (DIScUTIR MS). *Front Neurol*. 2017;8:65. doi:10.3389/fneur.2017.00065

**7**. Dijkstra IS, Pols J, Remmelts P, Brand PL. Preparedness for practice: a systematic cross-specialty evaluation of the alignment between postgraduate medical education and independent practice. *Med Teach*. 2015;37(2):153-161. doi:10.3109/0142159X.2014.929646

**8**. Phillips LS, Branch WT, Cook CB, et al. Clinical inertia. *Ann Intern Med*. 2001;135(9):825-834. doi:10.7326/0003-4819-135-9-200111060-00012

**9**. Okonofua EC, Simpson KN, Jesri A, Rehman SU, Durkalski VL, Egan BM. Therapeutic inertia is an impediment to achieving the Healthy People 2010 blood pressure control goals. *Hypertension*. 2006;47(3):345-351. doi:10.1161/01.HYP.0000200702.76436.4b

**10**. Khunti K, Gomes MB, Pocock S, et al. Therapeutic inertia in the treatment of hyperglycaemia in patients with type 2 diabetes: a systematic review. *Diabetes Obes Metab*. 2018;20(2):427-437. doi:10.1111/dom.13088

**11**. Lebeau JP, Cadwallader JS, Vaillant-Roussel H, et al. General practitioners' justifications for therapeutic inertia in cardiovascular prevention: an empirically grounded typology. *BMJ Open*. 2016;6(5):e010639. doi:10.1136/bmjopen-2015-010639

**12**. Nambam B, DuBose SN, Nathan BM, et al; T1D Exchange Clinic Network. Therapeutic inertia: underdiagnosed and undertreated hypertension in children participating in the T1D Exchange Clinic Registry. *Pediatr Diabetes*. 2016;17(1):15-20. doi:10.1111/pedi.12231

**13**. Edelson MG, Polania R, Ruff CC, Fehr E, Hare TA. Computational and neurobiological foundations of leadership decisions. *Science*. 2018;361(6401):eaat0036. doi:10.1126/science.aat0036

**14**. Lin H, Vartanian O. A neuroeconomic framework for creative cognition. *Perspect Psychol Sci*. 2018;13(6):655-677. doi:10.1177/1745691618794945

**15**. Aujoulat I, Jacquemin P, Rietzschel E, et al. Factors associated with clinical inertia: an integrative review. *Adv Med Educ Pract*. 2014;5:141-147. doi:10.2147/AMEP.S59022

**16**. Dohmen T, Falk A, Huffman D, Sunde U, Schupp J, Wagner GG. Individual risk attitudes: measurement, determinants, and behavioral consequences. *J Eur Econ Assoc*. 2011;9(3):522-550. doi:10.1111/j.1542-4774.2011.01015.x

**17**. Gerrity M, White K, DeVellis R, Dittus R. Physicians' reactions to uncertainty: refining the constructs and scales. *Motiv Emot*. 1995;19(3):175-191. doi:10.1007/BF02250510

**18**. Saposnik G, Sempere AP, Raptis R, Prefasi D, Selchen D, Maurino J. Decision making under uncertainty, therapeutic inertia, and physicians' risk preferences in the management of multiple sclerosis (DIScUTIR MS). *BMC Neurol*. 2016;16(1):58. doi:10.1186/s12883-016-0577-4

**19**. Prosperini L, Mancinelli CR, De Giglio L, De Angelis F, Barletta V, Pozzilli C. Interferon beta failure predicted by EMA criteria or isolated MRI activity in multiple sclerosis. *Mult Scler*. 2014;20(5):566-576. doi:10.1177/1352458513502399

**20**. Bermel RA, You X, Foulds P, et al. Predictors of long-term outcome in multiple sclerosis patients treated with interferon beta. *Ann Neurol*. 2013;73(1):95-103. doi:10.1002/ana.23758

**21**. Sormani MP, Gasperini C, Romeo M, et al; MAGNIMS Study Group. Assessing response to interferon-β in a multicenter dataset of patients with MS. *Neurology*. 2016;87(2):134-140. doi:10.1212/WNL.0000000000002830

**22**. Sormani MP, De Stefano N. Defining and scoring response to IFN-β in multiple sclerosis. *Nat Rev Neurol*. 2013;9(9):504-512. doi:10.1038/nrneurol.2013.146

**23**. Prosperini L, Giannì C, Leonardi L, et al. Escalation to natalizumab or switching among immunomodulators in relapsing multiple sclerosis. *Mult Scler*. 2012;18(1):64-71. doi:10.1177/1352458511417481

**24**. Saposnik G, Montalban X, Selchen D, et al. Therapeutic inertia in multiple sclerosis care: a study of Canadian neurologists. *Front Neurol*. 2018;9:781. doi:10.3389/fneur.2018.00781

**25**. Haneuse S, VanderWeele TJ, Arterburn D. Using the E-value to assess the potential effect of unmeasured confounding in observational studies. *JAMA*. 2019;321(6):602-603. doi:10.1001/jama.2018.21554

**26**. García Merino A, Ramón Ara Callizo J, Fernández Fernández O, Landete Pascual L, Moral Torres E, Rodríguez-Antigüedad Zarrantz A. Consensus statement on the treatment of multiple sclerosis by the Spanish Society of Neurology in 2016 [in Spanish]. *Neurologia*. 2017;32(2):113-119. doi:10.1016/j.nrl.2016.02.026

**27**. Cristiano E, Alonso R, Alvez Pinheiro A, et al. Argentinean recommendations on the identification of treatment failure in relapsing remitting multiple sclerosis patients. *J Neurol Sci*. 2018;385:217-224. doi:10.1016/j.jns.2018.01.004

**28**. Kaplan WA, Wirtz VJ, Stephens P. The market dynamics of generic medicines in the private sector of 19 low and middle income countries between 2001 and 2011: a descriptive time series analysis. *PLoS One*. 2013;8(9):e74399. doi:10.1371/journal.pone.0074399

**29**. Saposnik G, Maurino J, Sempere AP, et al. Overcoming therapeutic inertia in multiple sclerosis care: a pilot randomized trial applying the traffic light system in medical education. *Front Neurol*. 2017;8:430. doi:10.3389/fneur.2017.00430

**30**. Saposnik G, Mamdani M, Montalban X, et al. Traffic lights intervention reduces therapeutic inertia: a randomized controlled trial in multiple sclerosis care. *MDM P&P*. In press. doi:10.1177/2381468319855642

**31**. Regulatory process in the United States of America, Europe, China, and Japan. In: Shahzad A, ed. *Translational Medicine: Tools and Techniques*. Boston, MA: Academic Press; 2016:chap 5.

**32**. Garcia R, Araujo DV. The regulation of biosimilars in Latin America. *Curr Rheumatol Rep*. 2016;18(3):16. doi:10.1007/s11926-016-0564-1

**33**. Ernstsson O, Gyllensten H, Alexanderson K, Tinghög P, Friberg E, Norlund A. Cost of illness of multiple sclerosis: a systematic review. *PLoS One*. 2016;11(7):e0159129-e0159129. doi:10.1371/journal.pone.0159129

**34**. Hartung DM, Bourdette DN, Ahmed SM, Whitham RH. The cost of multiple sclerosis drugs in the US and the pharmaceutical industry: too big to fail? *Neurology*. 2015;84(21):2185-2192. doi:10.1212/WNL.0000000000001608

**35**. Saposnik G, Mamdani M, Terzaghi M, et al. The role of prescribing generic (non-proprietary) drugs in the prevalence of therapeutic inertia in multiple sclerosis care. *Front Neurol*. 2018;9:835. doi:10.3389/fneur.2018.00835

**SUPPLEMENT.**

**eAppendix.** Case Vignette Answered by Participating Physicians
**eFigure.** Observed vs Predicted TI Score After Adjustment for Covariates
**eTable 1.** Mixed Linear Regression Model Adjusted for Clustering
**eTable 2.** Bootstrap Linear Regression to Compare Normal vs Bias-Corrected 95% CI

# D. Appendix to Study 4

frontiers in Neurology

CLINICAL TRIAL
published: 21 August 2017
doi: 10.3389/fneur.2017.00430

## Overcoming Therapeutic Inertia in Multiple Sclerosis Care: A Pilot Randomized Trial Applying the Traffic Light System in Medical Education

*Gustavo Saposnik[1,2,3]*, Jorge Maurino[4], Angel P. Sempere[5], Maria A. Terzaghi[1,3], Christian C. Ruff[2], Muhammad Mamdani[6], Philippe N. Tobler[2†] and Xavier Montalban[1,7†]*

[1] Division of Neurology, Department of Medicine, St. Michael's Hospital, University of Toronto, Toronto, ON, Canada, [2] Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Zurich, Switzerland, [3] Outcomes Research and Decision Neuroscience Unit, Li Ka Shing Knowledge Institute, St. Michael's Hospital, University of Toronto, Toronto, ON, Canada, [4] Neuroscience Area, Medical Department, Roche Farma, Madrid, Spain, [5] Department of Neurology, Hospital General Universitario de Alicante, Alicante, Spain, [6] Healthcare Analytics Research, Li Ka Shing Knowledge Institute, St. Michael's Hospital, University of Toronto, Toronto, ON, Canada, [7] Neurology-Neuroimmunology Department, Neurorehabilitation Unit, Multiple Sclerosis Centre of Catalonia (Cemcat), Barcelona, Spain

OPEN ACCESS

**Edited by:**
*Ayrton R. Massaro, Hospital Sirio-Libanes, Brazil*

**Reviewed by:**
*Asif Doja, Children's Hospital of Eastern Ontario, Canada*
*Robert Jerome Frysztak, Loyola University Chicago, United States*

***Correspondence:**
*Gustavo Saposnik*
*saposnikg@smh.ca*

*†These authors have contributed equally to this work.*

**Specialty section:**
*This article was submitted to Neurology Education, a section of the journal Frontiers in Neurology*

**Received:** *06 July 2017*
**Accepted:** *08 August 2017*
**Published:** *21 August 2017*

**Citation:**
*Saposnik G, Maurino J, Sempere AP, Terzaghi MA, Ruff CC, Mamdani M, Tobler PN and Montalban X (2017) Overcoming Therapeutic Inertia*

**Background:** Physicians often do not initiate or intensify treatments when clearly warranted, a phenomenon known as therapeutic inertia (TI). Limited information is available on educational interventions to ameliorate knowledge-to-action gaps in TI.

**Objectives:** To evaluate the feasibility and efficacy of an educational intervention compared to usual care among practicing neurologists caring for patients with multiple sclerosis (MS).

**Methods:** We conducted a pilot double-blind, parallel-group, randomized clinical trial. Inclusion criteria included neurologists who are actively involved in managing MS patients. Participants were exposed to 20 simulated case-scenarios (10 cases at baseline, and 10 cases post-randomization to usual care vs. educational intervention) of relapsing–remitting MS with moderate or high risk of disease progression. The educational intervention employed a traffic light system (TLS) to facilitate decisions, allowing participants to easily recognize high-risk scenarios requiring treatment escalation. We also measured differences between blocks to invoke decision fatigue. The control group responded as they would do in their usual clinical practice not exposed to the educational intervention. The primary feasibility outcome was the proportion of participants who completed the study and the proportion of participants who correctly identified a high-risk case-scenario with the "red traffic light." Secondary outcomes included decision fatigue (defined as an increment of TI in the second block of case-scenarios compared to the first block) and the efficacy of the educational intervention measured as a reduction in TI for MS treatment.

This paper is published as Saposnik G, Mamdani M, Montalban X, Terzaghi M, Silva B, Saladino ML, Tobler PN, Caceres F. Traffic Lights Intervention Reduces Therapeutic Inertia: A Randomized Controlled Trial in Multiple Sclerosis Care. MDM Policy Pract., 4 (1): 2381468319855642, 2019.

frontiers in Neurology

CLINICAL TRIAL
published: 21 August 2017
doi: 10.3389/fneur.2017.00430

# Overcoming Therapeutic Inertia in Multiple Sclerosis Care: A Pilot Randomized Trial Applying the Traffic Light System in Medical Education

*Gustavo Saposnik[1,2,3]*, Jorge Maurino[4], Angel P. Sempere[5], Maria A. Terzaghi[1,3], Christian C. Ruff[2], Muhammad Mamdani[6], Philippe N. Tobler[2†] and Xavier Montalban[1,7†]*

[1] Division of Neurology, Department of Medicine, St. Michael's Hospital, University of Toronto, Toronto, ON, Canada, [2] Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Zurich, Switzerland, [3] Outcomes Research and Decision Neuroscience Unit, Li Ka Shing Knowledge Institute, St. Michael's Hospital, University of Toronto, Toronto, ON, Canada, [4] Neuroscience Area, Medical Department, Roche Farma, Madrid, Spain, [5] Department of Neurology, Hospital General Universitario de Alicante, Alicante, Spain, [6] Healthcare Analytics Research, Li Ka Shing Knowledge Institute, St. Michael's Hospital, University of Toronto, Toronto, ON, Canada, [7] Neurology-Neuroimmunology Department, Neurorehabilitation Unit, Multiple Sclerosis Centre of Catalonia (Cemcat), Barcelona, Spain

OPEN ACCESS

**Edited by:**
*Ayrton R. Massaro, Hospital Sirio-Libanes, Brazil*

**Reviewed by:**
*Asif Doja, Children's Hospital of Eastern Ontario, Canada*
*Robert Jerome Frysztak, Loyola University Chicago, United States*

***Correspondence:**
*Gustavo Saposnik*
*saposnikg@smh.ca*

*†These authors have contributed equally to this work.*

**Specialty section:**
*This article was submitted to Neurology Education, a section of the journal Frontiers in Neurology*

**Received:** *06 July 2017*
**Accepted:** *08 August 2017*
**Published:** *21 August 2017*

**Citation:**
*Saposnik G, Maurino J, Sempere AP, Terzaghi MA, Ruff CC, Mamdani M, Tobler PN and Montalban X (2017) Overcoming Therapeutic Inertia in Multiple Sclerosis Care: A Pilot Randomized Trial Applying the Traffic Light System in Medical Education. Front. Neurol. 8:430. doi: 10.3389/fneur.2017.00430*

**Background:** Physicians often do not initiate or intensify treatments when clearly warranted, a phenomenon known as therapeutic inertia (TI). Limited information is available on educational interventions to ameliorate knowledge-to-action gaps in TI.

**Objectives:** To evaluate the feasibility and efficacy of an educational intervention compared to usual care among practicing neurologists caring for patients with multiple sclerosis (MS).

**Methods:** We conducted a pilot double-blind, parallel-group, randomized clinical trial. Inclusion criteria included neurologists who are actively involved in managing MS patients. Participants were exposed to 20 simulated case-scenarios (10 cases at baseline, and 10 cases post-randomization to usual care vs. educational intervention) of relapsing–remitting MS with moderate or high risk of disease progression. The educational intervention employed a traffic light system (TLS) to facilitate decisions, allowing participants to easily recognize high-risk scenarios requiring treatment escalation. We also measured differences between blocks to invoke decision fatigue. The control group responded as they would do in their usual clinical practice not exposed to the educational intervention. The primary feasibility outcome was the proportion of participants who completed the study and the proportion of participants who correctly identified a high-risk case-scenario with the "red traffic light." Secondary outcomes included decision fatigue (defined as an increment of TI in the second block of case-scenarios compared to the first block) and the efficacy of the educational intervention measured as a reduction in TI for MS treatment.

**Results:** Of 30 neurologists invited to be part of the study, the participation rate was 83.3% ($n = 25$). Of the 25 participants, 14 were randomly assigned to the control group and 11 to the intervention group. TI was present in 72.0% of participants in at least one case scenario. For the primary feasibility outcome, the completion rate of the study

was 100% (25/25 participants). Overall, 77.4% of participants correctly identified the "red traffic light" for clinical-scenarios with high risk of disease progression. Similarly, 86.4% of participants correctly identified the "yellow traffic light" for cases that would require a reassessment within 6–12 months. For the secondary fatigue outcome, within-group analysis showed a significant increased prevalence of TI in the second block of case-scenarios (decision fatigue) among participants randomized to the control group (TI pre-intervention 57.1% vs. TI post-intervention 71.4%; $p = 0.015$), but not in the active group (TI pre-intervention 54.6% vs. TI post-intervention 63.6%; $p = 0.14$). For the efficacy outcome, we found a non-significant reduction in TI for the targeted intervention compared to controls (22.6 vs. 33.9% post-intervention; OR 0.57; 95% CI 0.26–1.22).

**Conclusion:** An educational intervention applying the TLS is feasible and shows some promising results in the identification of high-risk scenarios to reduce decision fatigue and TI. Larger studies are needed to determine the efficacy of the proposed educational intervention.

**Clinical Trial Registration:** www.ClinicalTrials.gov, identifier NCT03134794.

**Keywords: multiple sclerosis, disease-modifying therapy, neuroeconomics, decision making, risk aversion**

## BACKGROUND

Despite significant therapeutic advances, many patients remain undertreated, especially those with chronic medical conditions, such as atrial fibrillation, hypertension, and multiple sclerosis (MS) (1–4). One of the explanations relates to knowledge integration and knowledge-to-action gaps in therapeutic decisions. For example, it is known that physicians can be aware and informed about the current management of the common medical conditions they see in their daily clinical practice, but fail to integrate available information (e.g., severity of the condition, risk of progression, imaging findings, demographic factors affecting outcomes) and to implement best practice recommendations based on the available knowledge. This phenomenon may lead to therapeutic inertia (TI) usually associated with poorer outcomes (2–4). TI is a term that defines the absence of treatment initiation or intensification in patients when treatment goals are unmet. It affects 30–70% of clinicians caring for patients with chronic conditions (2, 5–7). Physician factors (e.g., low tolerance to uncertainty, *status quo* bias) are considered to be the main contributors to TI, but remain poorly studied (8–10).

Given physicians' limited training in risk management and formal learning in medical decision-making, educational interventions could optimize medical decisions (11). Previous research suggests that such interventions can improve medical decisions. A meta-analysis comprising 609 eligible studies enrolling 35,226 trainees compared the efficacy of simulation-based educational interventions (e.g., case-scenarios) in clinical skills and medical decision-making. The authors showed that a simulation-based educational intervention was more effective than standard educational programs for outcomes of knowledge, skills, and trainee's behavior (12). Other studies using a simulation-based intervention and clinical reasoning revealed a reduction in medical errors (13, 14). We have scarce information on strategies to overcome TI and only little evidence is available regarding effective educational interventions to reduce "knowledge-to-action" gaps.

The traffic light system (TLS) is an strategy that facilitates the decision-making process using traffic light terminology to match three types of situations: red light ("high risk"/"stop and think"), yellow light (warning), and green light ("stable"/"continue the same strategy"). The TLS emerged as a warning and risk categorization strategy to reduce human errors (15). It relies in a "hard-wired" cross-cultural color-coded concept that facilitates the integration of specific situations with an action (16–18). For example, studies showed that the TLS facilitated healthier food choices by interfering with automatic decisions and triggering re-evaluation processes (19). We focus on MS because of the broad availability of therapeutic options and clear definitions (clinical and radiological) of disease activity as the accepted criteria to escalate treatment.

We hypothesized that an educational intervention using the TLS may be feasible and effective to overcome insufficient knowledge integration and knowledge-to action gaps in the management of MS. In the present study, we evaluated the feasibility of an educational intervention to identify clinical situations of moderate and high risk of disease progression that may lead to TI. Our intervention was designed following the results of our previous studies on TI in MS care (4, 20).

## METHODS

### Study Design and Participants

This pilot, double-blinded, parallel-group, randomized clinical trial evaluated the feasibility of an educational intervention (active group) compared to usual care (control group) in the management of MS (**Figure 1**—CONSORT flow diagram). The goal of the education intervention was to facilitate risk stratification-action

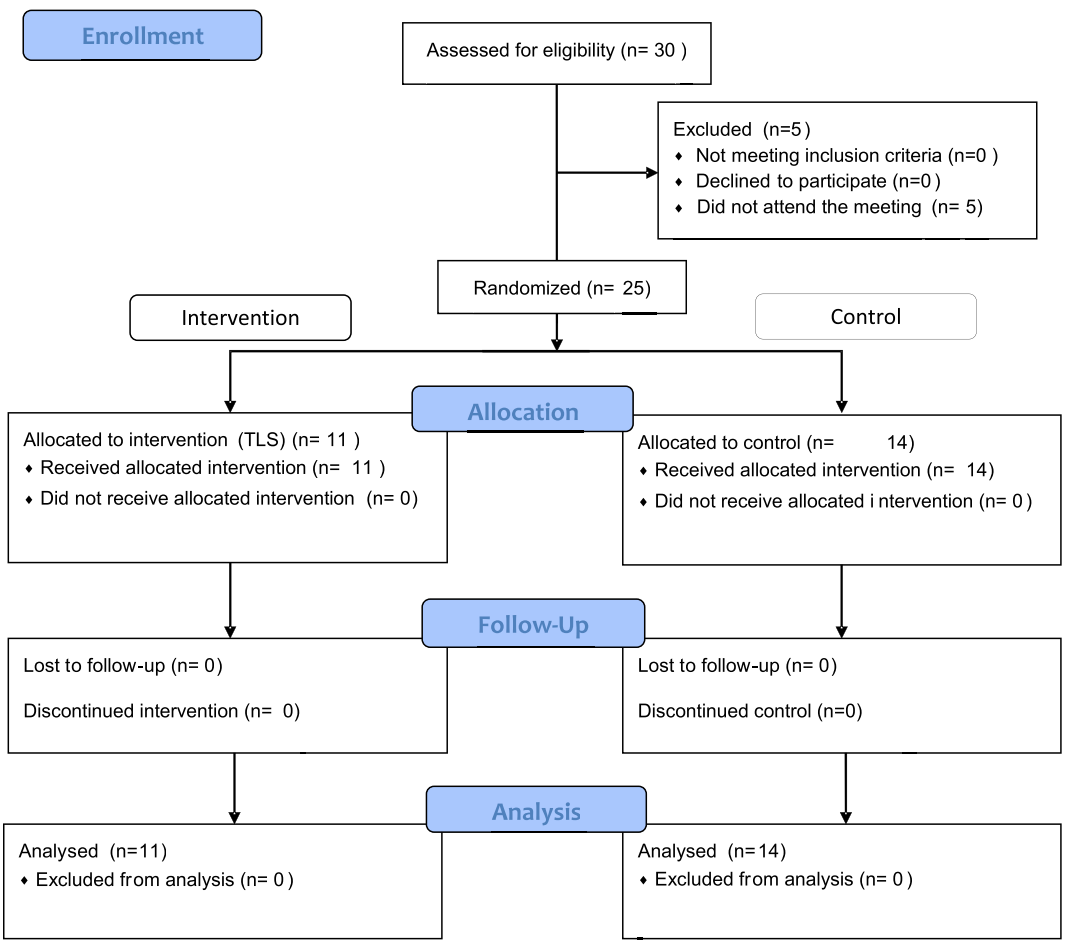



**FIGURE 1** | CONSORT flow diagram.

gaps in MS care. We expected the TLS would facilitate the identification of high-risk clinical-scenarios (i.e., "red" in the TLS) leading to assertive therapeutic decisions (e.g., escalate therapy when appropriate). Inclusion criteria included neurologists who were actively involved in managing MS patients. Physicians whose practice was primarily in caring for MS patients were classified as "MS specialists."

Candidates were invited to participate in a face-to-face meeting held in Madrid, Spain. The recruitment of participants was facilitated by the Spanish Neurological Society. We targeted the first 30 participants who replied to an e-mail invitation from a pool of neurologists who met the inclusion criteria. The study was conducted using Qualtrics, a web-based platform (www.Qualtrics.com). Each participant was provided with a tablet PC to complete the study. Participants were randomized (1:1 ratio), an automatic process in Qualtrics. Allocation concealment was facilitated in Qualtrics, so participants did not know what intervention will be allocated to after completing the 10 initial case-scenarios.

The study comprised 20 MS case-scenarios (see Appendix). Participants were exposed to 10 baselines case-scenarios (Block 1). Then, participants were randomized to usual care vs. educational intervention (TLS) followed by 10 additional similar case-scenarios (Block 2). In-line with the learning and education literature, case-vignettes, clinical scenarios, or "real world" encounters are regarded as the best simple strategy to evaluate cognitive biases among physicians (21, 22).

Case-scenarios were designed by our research team and MS experts (Angel P. Sempere, Gustavo Saposnik, Jorge Maurino, and Xavier Montalban). Overall, 16 cases were designed to assess appropriate escalation of treatment (absence of treatment escalation corresponding to TI; cases # 1–5, 8–10, 11–15, and 18–20), whereas the remaining four cases (case # 6, 7, 16, and 17) were designed to assess overtreatment (treatment escalation when there was no evidence of disease activity). Participants randomized to the intervention group (TLS) were also asked to identify the appropriate traffic light that would match the case-scenario. That question was prior to the selection of the therapeutic option.

In Block 2, eight cases corresponded to clinical situations of high risk of MS progression, which participants in the intervention group should associate with the "red traffic light," whereas two cases were associated with moderate risk of progression requiring a re-assessment in a 6- to 12-month period, which participants should associate with the "yellow traffic light" (see **Figures 2** and **3**).

Data management, research coordination, and statistical analyses were conducted at the Li Ka Shing Knowledge Institute of St. Michael's Hospital, Toronto. Operational procedures, guidelines for the implementation of both arms of the study, and the consent form were approved by the ethics review board at St. Michael's Hospital, University of Toronto. Online informed consent was obtained from all participants. The study was conducted in Spanish. Participants received compensation for transportation. Further details of the protocol were published in www.ClinicalTrials.gov # NCT03134794 and elsewhere (20).

## Rationale and Description of the Interventions

The current evidence suggests that medical decisions leading to TI are likely related to knowledge-to-action gaps (23, 24). We developed a simulation-based educational intervention aimed

| | |
|---|---|
| **1. INTERVENTION:** Provide a brief description of the educational intervention for all groups involved [e.g. control and comparator(s)]. | This educational intervention targets physicians involved in the care of patients with multiple sclerosis. It applies the TLS that allows learners to associate a color (red, yellow or green) with a medical situation facilitating therapeutic decisions. It introduced hard-wired concepts from Neuroeconomics (see figure 3). |
| WHY -this educational process | |
| **2. THEORY:** Describe the educational theory (ies), concept or approach used in the intervention. | The current evidence suggests that medical decisions leading to TI are likely related to knowledge-to-action gaps rather than knowledge gaps. The action component is the therapeutic decision (e.g. continue on the same treatment, change for another one that would NOT affect the clinical course or escalating to a more effective agent). A simulation-based educational intervention facilitates the integration of relevant concepts for therapeutic decisions (e.g. disease course and severity, risk stratification, and therapeutic options). Specifically, the TLS, an educational intervention, facilitates the integration of medical knowledge and therapeutic decisions. Previous studies using neuroeconomic methods showed that a color-coded system influences the valuation process (mechanism by which our brain ranks different options) facilitating optimal choices. fMRI studies showed that TLS labels enhance the coupling between brain regions associated with valuation (vmPFC) and self-control. Specifically, a "red" traffic light activated the left inferior frontal gyrus/dlPFC, a region implicated in self-control in food choices and tolerance to uncertainty. The TLS applied to medical care facilitates the distinction of clinical situations that require a proactive decision (e.g. escalate therapy due to a high risk of poor outcomes), possibly by ameliorating physicians' low tolerance to uncertainty. |
| **3. LEARNING OBJECTIVES:** Describe the learning objectives for all groups involved in the educational intervention. | 1) To reduce TI in clinical care.<br>2) To facilitate the integration of knowledge with medical decisions to better evidence based practice. |
| **4. EVIDENCE BASED PRACTICE CONTENT:** List the foundation steps of EBP (ask, acquire, appraise, apply, assess) included in the educational intervention. | Ask about the current clinical scenario. Integrates the evidence for treatment escalation. Appraise the expected outcome given clinical and imaging activity. Evaluates the selection of the traffic light that corresponds to the case-scenario to guide therapeutic decisions. The TLS is a simple strategy that can be easily disseminated among treating physicians. |
| WHAT<br>**5. MATERIALS:** Describe the specific educational materials used in the educational intervention. Include materials provided to the learners and those used in the training of educational intervention providers | Our educational intervention integrates current knowledge in MS care with the TLS (see figure). Participants are provided with the following statements:<br>1) MS care is becoming more complex (e.g.: increasing therapeutic options, new paradigms: injectable, oral, and infusion therapies)<br>2) Despite the recent advances, many MS patients remain undertreated according to the best clinical practice guidelines.<br>3) As physicians, we have received limited training in risk management and formal education in decision making.<br>4) The TLS emerged as a strategy to help optimize/facilitate choices and decisions. TLS applied to food labels successfully improved healthy choices. |

**FIGURE 2** | Continued

at facilitating the integration of knowledge and overcoming knowledge-to-action gaps in MS care (25). The *action* component is the *therapeutic decision* (e.g., continue on the same treatment, change to a treatment that would not affect the clinical course or escalating to a more effective agent). We followed the Guideline for Reporting of Evidence-based practice Educational interventions and Teaching (GREET) statement to describe our educational intervention (**Figure 2**) (26).

## Educational Intervention: The TLS

Our study included two phases: pre-intervention and post-intervention periods (**Figure 3**). Participants were randomly assigned to the educational intervention or control groups after the pre-intervention period.

Our educational intervention is based on the application of the TLS to medical decision-making (16–19, 27). In our study, the TLS was applied to help participants identify high-risk

| | |
|---|---|
| | 5) Clinicians can apply their knowledge using the TLS to improve therapeutic decisions: |
| **6. EDUCATIONAL STRATEGIES:** Describe the teaching/learning strategies (e.g. tutorials, lectures, online modules) used in the educational intervention. | This is based on an online module (See figure 3) integrated with simulated case-scenarios. |
| **7. INCENTIVES:** Describe any incentives or reimbursements provided to the learners. | Learners are provided with an incentive for research studies or reimbursement for travelling when participating in educative workshops. |
| WHO PROVIDED | This is based on an online module integrated with simulated case-scenarios. No instructors are needed. |
| **8. INSTRUCTORS:** For each instructor(s) involved in the educational intervention describe their professional discipline, teaching experience/expertise. Include any specific training related to the educational intervention provided for the instructor(s). | |
| HOW | Online module integrated with simulated case-scenarios. It could also be applied in face-to-face meetings/workshops using tablet PC or slides to explain the intervention. |
| **9. DELIVERY:** Describe the modes of delivery (e.g. face-to-face, internet or independent study package) of the educational intervention. Include whether the intervention was provided individually or in a group and the ratio of learners to instructors. | |
| WHERE | This educational intervention is simple and can be delivered in any learning space with an electronic device and internet, including remote participation from home. |
| **10. ENVIRONMENT:** Describe the relevant physical learning spaces (e.g. conference, university lecture theatre, hospital ward, community) where the teaching/learning occurred. | |
| WHEN and HOW MUCH | This is a simple self-learning educational intervention requiring a single session that takes approximately 5 to 10 minutes. |
| **11. SCHEDULE:** Describe the scheduling of the educational intervention including the number of sessions, their frequency, timing and duration. | |
| **12.** Describe the amount of time learners spent in face to face contact with instructors and any designated time spent in self-directed learning activities. | |
| **PLANNED CHANGES** | No adaptation other than translation to different languages |

**FIGURE 2** | Continued

| | |
|---|---|
| **13.** Did the educational intervention require specific adaptation for the learners? If yes, please describe the adaptations made for the learner(s) or group(s). | |
| **UNPLANNED CHANGES** | No changes during the study |
| **14.** Was the educational intervention modified during the course of the study? If yes, describe the changes (what, why, when, and how). | |
| HOW WELL | 30 participants were invited. 25 attended our workshop on decision making and TI in MS care. All 25 participants completed the study. Each participant received a tablet to complete the study. There was no deviation from the protocol. The study consisted of three parts: i) risk preferences, ii) Simulated case-scenarios I, iii) Simulated case-scenarios II (after participants were randomly assigned to the educational intervention or standard practice). |
| **15. ATTENDANCE:** Describe the learner attendance, including how this was assessed and by whom. Describe any strategies that were used to facilitate attendance. | |
| **16.** Describe any processes used to determine whether the materials (item 5) and the educational strategies (item 6) used in the educational intervention were delivered as originally planned. | |
| **17.** Describe the extent to which the number of sessions, their frequency, timing and duration for the educational intervention were delivered as scheduled (item 11). | |

**FIGURE 2** | Description of the educational intervention according to the GREET guidelines.

cases-scenarios, where MS patients had both clinical and radiological activity. Consequently, participants should be able to identify the "red" traffic light and escalate treatment. The "yellow" represents caution when MS patients had either a clinical relapse or some degree of activity on brain imaging (but not both), which requires a reassessment within 6–12 months.

The control group made therapeutic decisions without being exposed to the educational intervention as part of the current standard practice. They had the option to take a break or continue the study. The estimated time of study completion per participant ranged between 30 and 35 min and did not differ between groups.

### Definitions

For the primary analysis, high risk of progression was defined as the combination of a clinical relapse plus the presence of new brain lesions in follow-up magnetic resonance imaging (MRI) scans or at least one gadolinium-enhancing lesion (28, 29). All high-risk simulated clinical cases included a description of an MRI with more than five new T2 lesions or at least one enhancing lesion (30). The use of these definitions combining a clinical relapse and MRI activity is consistent with recent evidence regarding the risk of treatment failure among patients receiving interferon-β (31). Disease progression was defined as at least one point worsening from baseline in the Expanded Disability Status Scale score (32).

Recent meta-analysis confirmed that alemtuzumab, natalizumab, and fingolimod are the best available choices for preventing clinical relapses in patients with relapsing–remitting MS (RRMS) (33). The current treatment option for RRMS include first-line (beta interferons, glatiramer acetate), second-line (fingolimod), and third-line (natalizumab, alemtuzumab) therapies. For the present analysis, we used the aforementioned scheme according to the current clinical practice (4, 34, 35).

## Outcome Measures

The primary feasibility outcome was the proportion of participants who completed the study. A completion rate of 70% or higher was our pre-specified outcome. The feasibility of delivering the intervention was defined as the number of participants who correctly identified the "red traffic light" for clinical-scenarios comprising a high-risk of progression. A pre-specified criterion of at least 70% correct responses was used to classify the educational intervention as feasible.

Efficacy of the educational intervention, a secondary outcome measure, was defined as a reduction in TI based on each individual response. We also evaluated secondary outcome the capability of the intervention to protect against decision fatigue decision fatigue [defined as the difference in TI within groups before (Block 1) and after the intervention (Block 2)] (36, 37). A significantly higher prevalence of TI in the 10 case-scenarios post-intervention (Block 2) would be indicative of decision fatigue.

## Statistical Analysis

Given the pilot nature of this study, we performed primarily descriptive statistics. We used non-parametric tests to compare continuous and categorical variables between groups. A Welch's *t*-test was used to rule out large differences in TI. Logistic regression analysis was completed to determine the efficacy of the educational intervention in the reduction of TI after adjusting for responses in the pre-intervention period (Block 1).

We evaluated two different outcome measures: (i) TI defined as lack of treatment escalation in at least one case scenario and (ii) number of participants' responses representing TI. We also compared the total number of correct responses for each case-scenario between and within groups before and after the intervention.

Given our pre-specified target intervention, we compared the proportion of responses associated with TI between those who selected the "red light" in the active group vs. control group.

All tests were two-tailed, and $p$-values <0.05 were considered significant. We used STATA 13 (College Station, TX, USA: StataCorp LP) to conduct all analyses.

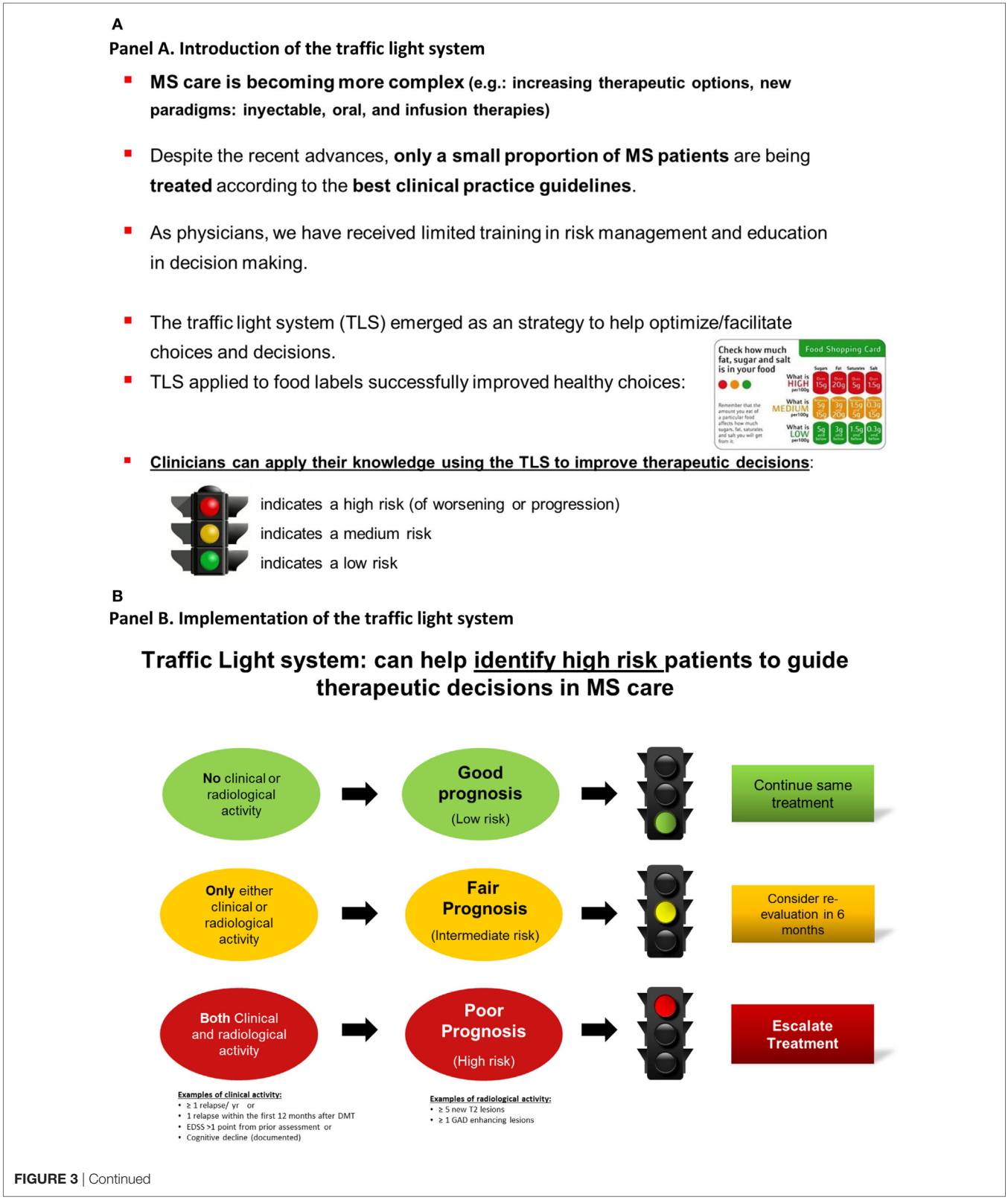



**FIGURE 3** | Continued

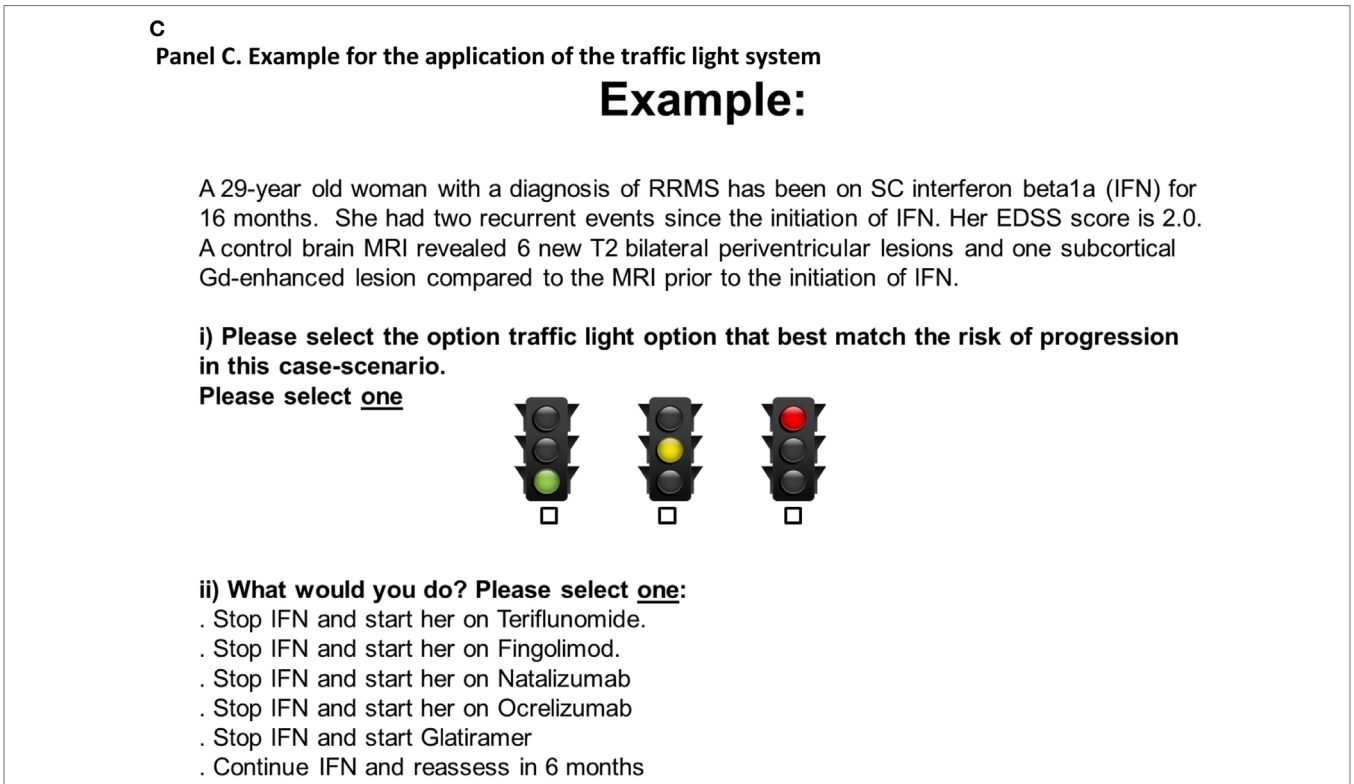



**FIGURE 3** | Educational intervention: the traffic light system may facilitate therapeutic decisions in multiple sclerosis care. Participants viewed the two informative panels **(A,B)** and a third panel providing an example **(C)**.

## RESULTS

Of 30 neurologists from across Spain who were invited to participate in the study, 25 (83.3%) attended the meeting. Eleven participants were randomly assigned to the educational intervention, whereas the remaining 14 were assigned to the control group.

Overall, the mean (SD) age was 35.4 (±7.3) years; 16 (64%) were females. Sixty percent (15/25) of participants primarily focused their practice on MS care. **Table 1** summarizes baseline characteristics of the study population. Baseline characteristics appeared similar between groups. None of the participants choose to have a break.

The mean time to start the second set of cases post-randomization (block 2) was 12.7 s in the control group and 11.8 s in the intervention group.

For the primary feasibility outcome, the completion rate of the study was 100% (25/25 participants). TI was present in 72.0% of participants in at least one case scenario. Only 4 (16%) participants did not exhibit TI (all case-scenarios were correct), whereas one-third of participants (8/25) exhibited TI in five or more simulated case-scenarios. In the within-group analysis, we observed an increased prevalence of TI in the second set of case-scenarios (defined as decision fatigue) among participants in the control group (TI pre-intervention 57.1% vs. TI post-intervention 71.4%; $p = 0.015$), but not in the active group (TI pre-intervention 54.6% vs. TI post-intervention 63.6%; $p = 0.14$). Decision fatigue was associated with higher odds of TI (OR 3.99; 95% CI 1.05–15.1).

**TABLE 1** | Baseline characteristics of participants.

| Characteristics | Total (%) $n = 25$ | Intervention (%) $n = 11$ | Control (%) $n = 14$ |
|---|---|---|---|
| Age (mean ± SD), in years | 35.4 ± 7.3 | 33.8 ± 5.5 | 36.6 ± 8.4 |
| Sex | | | |
| Female | 16 (64.0) | 8 (72.7) | 8 (57.1) |
| Specialty | | | |
| Multiple sclerosis (MS) specialists | 15 (60.0) | 6 (54.6) | 9 (64.3) |
| General neurologists who care for MS patients | 10 (40.0) | 5 (45.5) | 5 (35.7) |
| Practice setting | | | |
| Academic | 22 (88.0) | 11 (100) | 11 (78.6) |
| Community | 2 (8.0) | 0 (0) | 2 (14.3) |
| Both (academic and non-academic) | 1 (4.0) | 0 (0) | 1 (7.1) |
| % time in clinical practice | | | |
| Greater than 75% | 16 (64.0) | 6 (54.6) | 10 (71.4) |
| Years in practice, mean (±SD) | 9.9 ± 7.3 | 8.5 ± 5.0 | 11.1 ± 8.7 |
| MS patients seen per week, mean (±SD) | 17 ± 11.0 | 15.9 ± 10.5 | 17.8 ± 11.7 |
| Attended latest ECTRIMS conference | 14 (56.0) | 5 (45.5) | 9 (64.3) |
| Author of a peer-reviewed publication in the last 12 months | 12 (48.0) | 3 (27.3) | 9 (64.3) |

*Numbers in brackets indicate percentages, unless otherwise indicated.*

There was no TI block-by-intervention group interaction ($p = 0.61$). Comparative results between pre- and post-intervention within and between groups are summarized in **Table 2**.

**TABLE 2** | Efficacy outcome measures: comparison pre- and post-intervention within and between groups.

| Efficacy outcomes | Intervention group | | | Control group | | | Comparison between groups |
|---|---|---|---|---|---|---|---|
| | Pre-intervention | Post-intervention | Change from pre-intervention (95% CI)[a] | Pre-intervention | Post-intervention | Change from pre-intervention (95% CI)[a] | Differences between groups post-intervention (95% CI)[b] |
| | *n* = 11 | *n* = 11 | | *n* = 14 | *n* = 14 | | |
| Mean (SD) number of individual responses related to therapeutic inertia (TI) | 1.91 (1.3) | 2.36 (1.5) | 0.45 (–2.20, 1.31) | 2.14 (1.4) | 2.71 (1.8) | 0.57 (–2.61, 1.48) | 0.35 (–1.01, 1.71) |
| Individual responses related to TI, *n/N* (%) | 21/88 (0.24) | 26/88 (0.30) | 0.056 (–1.30, 1.41) | 30/112 (0.27) | 38/112 (0.34) | 0.071 (–1.18, 1.32) | 0.044 (–1.33, 1.42) |
| % (SD) of participants with TI[c] | 54.5 (27.2) | 63.6 (25.4) | 9.1 (–14.2, 32.4) | 57.1 (26.4) | 71.4 (22.0) | 14.3 (–4.5, 33.1) | 7.80 (–12.2, 27.8) |

*n, number of responses related with TI; N, total number of responses. Numbers were rounded to two decimals.*

[a]*Represents the difference and 95% CI in the efficacy outcomes between pre- and post-intervention within groups.*

[b]*Represents the difference and 95% CI in the efficacy outcomes post-intervention between groups.*

[c]*TI identified in at least one case-scenario.*

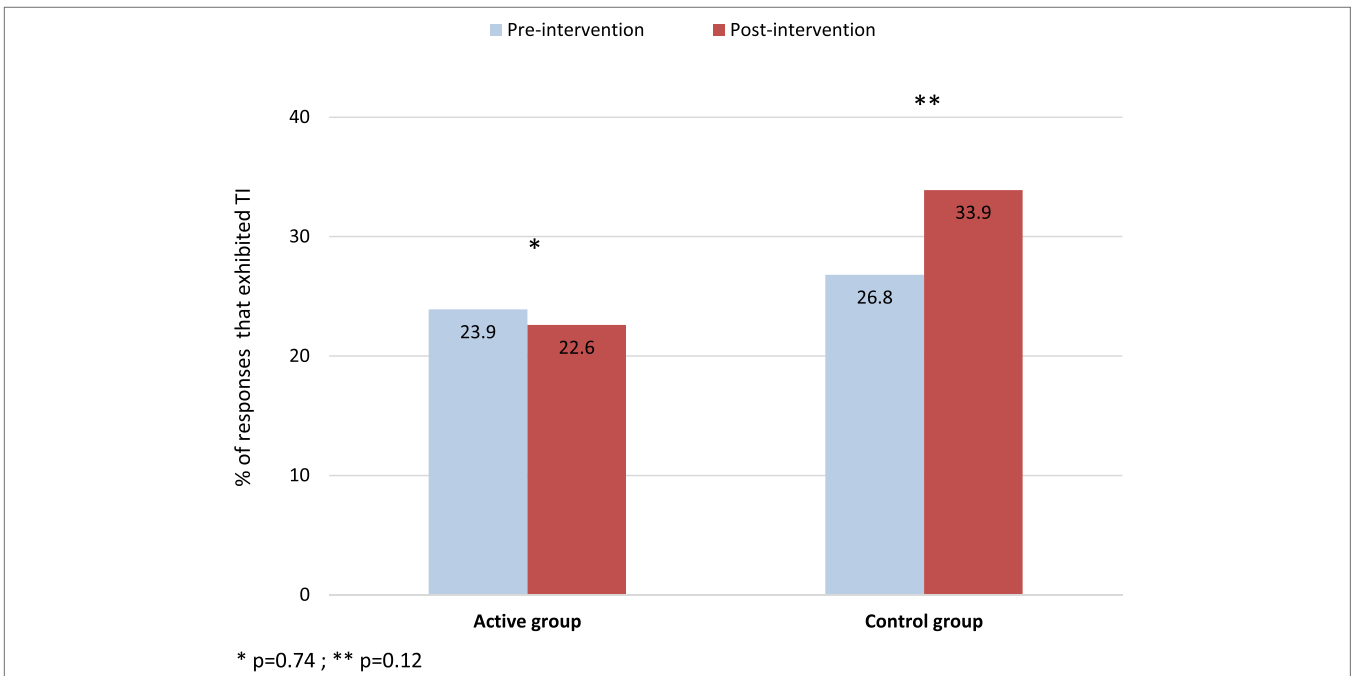



**FIGURE 4** | Prevalence of therapeutic inertia (TI) for the targeted intervention accounting for each individual response in the active and control groups. Lower numbers represent lower TI (more optimal therapeutic decisions). *$p = 0.74$, **$p = 0.12$. Note the lower trend in the prevalence of TI in the intervention group compared to the control group (22.6 vs. 33.9%; OR 0.57; 95% CI 0.26–1.22).

Overall, 77.4% of participants correctly identified the "red traffic light" for clinical scenarios with high-risk of disease progression. Similarly, 86.4% of participants correctly identified the "yellow traffic light" for cases that would require a reassessment within 6–12 months. Thus, participants knew what should be done with different cases.

The analysis of each individual case-scenario revealed that TI was present in 23.9% of responses in the interventional group and 26.8% of responses in the control group in the pre-intervention period ($p = 0.74$) (**Figure 4**). The multivariate analysis of each individual response revealed a non-significant reduction in TI in favor of the intervention group (OR 0.82; 95% CI 0.25–2.69) after adjusting for pre-intervention TI. The analysis evaluating individual responses targeted by the intervention (those cases where participants correctly identified the red light for high-risk scenarios) revealed a non-significant reduction of TI in the

157

intervention group compared to the control group (OR 0.57; 95% CI 0.26–1.22) (**Figure 4**).

## DISCUSSION

Therapeutic inertia is a common phenomenon in the management of patients with chronic medical conditions (2, 8). Moreover, there is limited information regarding educational interventions to overcome the effects of TI. In the present study, we used the well-defined paradigm of MS care, with its broad variability of therapeutic options to escalate treatment as a response to evidence in disease activity.

We conducted a pilot randomized study allocating participating neurologists to an educational intervention (using the TLS) or usual care (control group). We found that TI was present in at least one case-scenario in 7 out of 10 participants. Overall, the great majority of participants correctly matched the traffic light (yellow and red) with the simulated case-scenario and appropriately escalated treatment. We found a non-significant 43% reduction in the odds of TI by identifying the red traffic light. We also identified decision fatigue in the control group, but not in the intervention group. This finding suggests that the educational intervention may promote the continuity of accurate therapeutic decisions over time (by ameliorating the impact of decision fatigue on TI) despite the increasing number of case-scenarios.

The use of the TLS is a novel initiative to optimize decisions. It has been successfully applied to different medical fields, including the selection of healthier food choices leading to weight loss or the detection of children with fever at high risk of developing a serious bacterial infection (16, 18). Stangel and colleagues proposed the TLS to monitor treatment response in patients with RRMS. They included a more sophisticated scoring system (0–3) to categorize clinical relapses, evidence of disease progression, a cognitive assessment, and MRI findings. This scoring system leads to a decision model that uses the TLS to facilitate therapeutic choices (38). At the time of writing this manuscript, there were no data available on the application of this strategy.

Our study has some significant limitations. First, the sample size is small given the pilot design. As a consequence, our study was not powered to determine the efficacy of the educational intervention. Second, we used simulated case-scenarios that may not accurately reflect therapeutic decisions in clinical practice or known patients followed up over time. Third, some participants' responses may reflect local limitations in the prescription of disease modifying agents. Fourth, we only tested some physician-level factors that may influence TI. Finally, we do not know if the educational intervention would require reinforcement months later to maintain its potential effect on TI.

Despite these limitations, our study suggests that a simple educational intervention applying the TLS is feasible to increase clinician's recognition of MS patients at high risk of progression and overcome decision fatigue. Although our study evaluated therapeutic decisions in MS, the educational intervention could be applied to the management of other medical conditions and thus have wider-reaching implications for clinical care. Furthermore, our results serve as the basis for sample size calculations in the design of future studies.

Increasing awareness is the first step in the decision-making process to reduce the effects of TI. We used the TLS to increase awareness of treatment-relevant knowledge. Our study is also strengthened by: (i) a randomized design, (ii) the application of an evidence-based educational approach following the GREET guidelines (26), (iii) the implementation of a simple educational intervention that links to the neural pathways involved in decision-making under uncertainty (19, 27), and (iv) the target of a clinically relevant outcome (i.e., TI and decision fatigue) (4, 36) with the goal of overcoming knowledge-to-action gaps in MS treatment.

The next steps would include the implementation of studies at a larger scale to determine the efficacy of our educational intervention in overcoming decision fatigue and reducing TI among primary care physicians and specialists managing patients with neurological (MS, stroke) and other chronic conditions (e.g., atrial fibrillation, diabetes).

## ETHICS STATEMENT

This study was carried out in accordance with the recommendations of St. Michael's Hospital Ethical Committee with online informed consent from all subjects. All subjects gave online informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the St. Michael's Hospital Ethical Committee.

## AUTHOR CONTRIBUTIONS

GS: study concept and design, creation of the educational intervention, acquisition of data, analysis and interpretation of the data, and obtaining funding. AS: study concept and design, interpretation of the data, and critical revision of the manuscript for intellectual content. MT: study implementation, interpretation of the data, and critical revision of the manuscript for intellectual content. MM: study concept and design, design of the educational intervention, interpretation of the data, and critical revision of the manuscript for intellectual content. CR, JM, and PT: study concept and design, interpretation of the data, critical revision of the manuscript for intellectual content, and study supervision. XM: supervision of MS case-scenarios, interpretation of the data, and critical revision of the manuscript for intellectual content.

## ACKNOWLEDGMENTS

The authors are most grateful to all physicians participating in the study. We appreciate the support from the Department of Economics at the University of Zurich by facilitating access to Qualtrics platform. We also appreciate the support of Dr. Ryan

Brydges and Nikki Woods of the Wilson Center for Medical Education at the University of Toronto.

## FUNDING

The study was sponsored by the Sociedad Española de Neurologia (SEN) and funded by an operating grant from Roche Farma Spain. The sponsors were not involved in the design, execution, analysis, and interpretation or reporting of the results.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://journal.frontiersin.org/article/10.3389/fneur.2017.00430/full#supplementary-material.

## REFERENCES


1. Nieuwlaat R, Olsson SB, Lip GY, Camm AJ, Breithardt G, Capucci A, et al. Guideline-adherent antithrombotic treatment is associated with improved outcomes compared with undertreatment in high-risk patients with atrial fibrillation. The Euro heart survey on atrial fibrillation. *Am Heart J* (2007) 153(6):1006–12. doi:10.1016/j.ahj.2007.03.008
2. O'Connor PJ, Sperl-Hillen JAM, Johnson PE, Rush WA, Biltz G. Clinical inertia and outpatient medical errors. In: Henriksen K, Battles JB, Marks ES, Lewin DI, editors. *Advances in Patient Safety: From Research to Implementation*. (Vol. 2), Rockville, MD: Concepts and Methodology (2005).
3. Ziemssen T, Derfuss T, de Stefano N, Giovannoni G, Palavra F, Tomic D, et al. Optimizing treatment success in multiple sclerosis. *J Neurol* (2016) 263(6):1053–65. doi:10.1007/s00415-015-7986-y
4. Saposnik G, Sempere AP, Prefasi D, Selchen D, Ruff CC, Maurino J, et al. Decision-making in multiple sclerosis: the role of aversion to ambiguity for therapeutic inertia among neurologists (DIScUTIR MS). *Front Neurol* (2017) 8:65. doi:10.3389/fneur.2017.00065
5. Mohan AV, Phillips LS. Clinical inertia and uncertainty in medicine. *JAMA* (2011) 306(4):383; author reply–4. doi:10.1001/jama.2011.1045
6. Phillips LS, Branch WT, Cook CB, Doyle JP, El-Kebbi IM, Gallina DL, et al. Clinical inertia. *Ann Intern Med* (2001) 135(9):825–34. doi:10.7326/0003-4819-135-9-200111060-00012
7. Okonofua EC, Simpson KN, Jesri A, Rehman SU, Durkalski VL, Egan BM. Therapeutic inertia is an impediment to achieving the healthy people 2010 blood pressure control goals. *Hypertension* (2006) 47(3):345–51. doi:10.1161/01.HYP.0000200702.76436.4b
8. Lebeau JP, Cadwallader JS, Aubin-Auger I, Mercier A, Pasquet T, Rusch E, et al. The concept and definition of therapeutic inertia in hypertension in primary care: a qualitative systematic review. *BMC Fam Pract* (2014) 15:130. doi:10.1186/1471-2296-15-130
9. Lebeau JP, Cadwallader JS, Vaillant-Roussel H, Pouchain D, Yaouanc V, Aubin-Auger I, et al. General practitioners' justifications for therapeutic inertia in cardiovascular prevention: an empirically grounded typology. *BMJ Open* (2016) 6(5):e010639. doi:10.1136/bmjopen-2015-010639
10. Simon D. Therapeutic inertia in type 2 diabetes: insights from the PANORAMA study in France. *Diabetes Metab* (2012) 38(Suppl 3):S47–52. doi:10.1016/S1262-3636(12)71534-3
11. Dijkstra IS, Pols J, Remmelts P, Brand PL. Preparedness for practice: a systematic cross-specialty evaluation of the alignment between postgraduate medical education and independent practice. *Med Teach* (2015) 37(2):153–61. doi:10.3109/0142159X.2014.929646
12. Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. *JAMA* (2011) 306(9):978–88. doi:10.1001/jama.2011.1234
13. Khunti K, Millar-Jones D. Clinical inertia to insulin initiation and intensification in the UK: a focused literature review. *Prim Care Diabetes* (2017) 11:3–12. doi:10.1016/j.pcd.2016.09.003
14. Cooke CE, Sidel M, Belletti DA, Fuhlbrigge AL. Review: clinical inertia in the management of chronic obstructive pulmonary disease. *COPD* (2012) 9(1):73–80. doi:10.3109/15412555.2011.631957
15. Orriols L, Delorme B, Gadegbeku B, Tricotel A, Contrand B, Laumon B, et al. Prescription medicines and the risk of road traffic crashes: a French registry-based study. *PLoS Med* (2010) 7(11):e1000366. doi:10.1371/journal.pmed.1000366
16. Sonnenberg L, Gelsomin E, Levy DE, Riis J, Barraclough S, Thorndike AN. A traffic light food labeling intervention increases consumer awareness of health and healthy choices at the point-of-purchase. *Prev Med* (2013) 57(4):253–7. doi:10.1016/j.ypmed.2013.07.001
17. Hawley KL, Roberto CA, Bragg MA, Liu PJ, Schwartz MB, Brownell KD. The science on front-of-package food labels. *Public Health Nutr* (2013) 16(3):430–9. doi:10.1017/S1368980012000754
18. Murphy MS, Baker M. When used appropriately, NICE's traffic light system for febrile children helps clinicians make safe and appropriate decisions. *BMJ* (2014) 348:g2518. doi:10.1136/bmj.g2518
19. Enax L, Krajbich I, Weber B. Salient nutrition labels increase the integration of health attributes in food decision-making. *Judgment Decis Mak* (2016) 11(5):460–71.
20. Saposnik G, Sempere AP, Raptis R, Prefasi D, Selchen D, Maurino J. Decision making under uncertainty, therapeutic inertia, and physicians' risk preferences in the management of multiple sclerosis (DIScUTIR MS). *BMC Neurol* (2016) 16(1):58. doi:10.1186/s12883-016-0577-4
21. Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *JAMA* (2000) 283(13):1715–22. doi:10.1001/jama.283.13.1715
22. Peabody JW, Luck J, Glassman P, Jain S, Hansen J, Spell M, et al. Measuring the quality of physician practice by using clinical vignettes: a prospective validation study. *Ann Intern Med* (2004) 141(10):771–80. doi:10.7326/0003-4819-141-10-200411160-00008
23. Kohok DD, Shah S, Kumar R, Venuto L, Gudleski G, Venuto R. Interventions to reduce clinical inertia in cardiac risk factor management in renal transplant recipients. *J Clin Hypertens (Greenwich)* (2014) 16(2):127–32. doi:10.1111/jch.12249
24. Gongora-Ortega J, Segovia-Bernal Y, Valdivia-Martinez Jde J, Galaviz-DeAnda JM, Prado-Aguilar CA. Educational interventions to improve the effectiveness in clinical competence of general practitioners: problem-based versus critical reading-based learning. *BMC Med Educ* (2012) 12:53. doi:10.1186/1472-6920-12-53
25. Kennedy T, Regehr G, Rosenfield J, Roberts SW, Lingard L. Exploring the gap between knowledge and behavior: a qualitative study of clinician action following an educational intervention. *Acad Med* (2004) 79(5):386–93. doi:10.1097/00001888-200405000-00006
26. Phillips AC, Lewis LK, McEvoy MP, Galipeau J, Glasziou P, Moher D, et al. Development and validation of the guideline for reporting evidence-based practice educational interventions and teaching (GREET). *BMC Med Educ* (2016) 16(1):237. doi:10.1186/s12909-016-0759-1
27. Enax L, Hu Y, Trautner P, Weber B. Nutrition labels influence value computation of food products in the ventromedial prefrontal cortex. *Obesity (Silver Spring)* (2015) 23(4):786–92. doi:10.1002/oby.21027
28. Prosperini L, Mancinelli CR, De Giglio L, De Angelis F, Barletta V, Pozzilli C. Interferon beta failure predicted by EMA criteria or isolated MRI activity in multiple sclerosis. *Mult Scler* (2014) 20(5):566–76. doi:10.1177/1352458513502399
29. Bermel RA, You X, Foulds P, Hyde R, Simon JH, Fisher E, et al. Predictors of long-term outcome in multiple sclerosis patients treated with interferon beta. *Ann Neurol* (2013) 73(1):95–103. doi:10.1002/ana.23758
30. Sormani MP, Rio J, Tintore M, Signori A, Li D, Cornelisse P, et al. Scoring treatment response in patients with relapsing multiple sclerosis. *Mult Scler* (2013) 19(5):605–12. doi:10.1177/1352458512460605
31. Sormani MP, Gasperini C, Romeo M, Rio J, Calabrese M, Cocco E, et al. Assessing response to interferon-beta in a multicenter dataset of


159

patients with MS. *Neurology* (2016) 87(2):134–40. doi:10.1212/WNL.0000000000002830
32. Sormani MP, De Stefano N. Defining and scoring response to IFN-beta in multiple sclerosis. *Nat Rev Neurol* (2013) 9(9):504–12. doi:10.1038/nrneurol.2013.146
33. Tramacere I, Del Giovane C, Salanti G, D'Amico R, Filippini G. Immunomodulators and immunosuppressants for relapsing-remitting multiple sclerosis: a network meta-analysis. *Cochrane Database Syst Rev* (2015) 9:CD011381. doi:10.1002/14651858.CD011381.pub2
34. Prosperini L, Gianni C, Leonardi L, De Giglio L, Borriello G, Galgani S, et al. Escalation to natalizumab or switching among immunomodulators in relapsing multiple sclerosis. *Mult Scler* (2012) 18(1):64–71. doi:10.1177/1352458511417481
35. Garcia-Merino A, Fernandez O, Montalban X, de Andres C, Oreja-Guevara C, Rodriguez-Antiguedad A, et al. Consensus statement on medication use in multiple sclerosis by the Spanish Society of Neurology's study group for demyelinating diseases. *Neurologia* (2013) 28(6):375–8. doi:10.1016/j.nrl.2013.01.009
36. Ancker JS, Edwards A, Nosal S, Hauser D, Mauer E, Kaushal R, et al. Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. *BMC Med Inform Decis Mak* (2017) 17(1):36. doi:10.1186/s12911-017-0430-8
37. Stewart AF, Ferriero DM, Josephson SA, Lowenstein DH, Messing RO, Oksenberg JR, et al. Fighting decision fatigue. *Ann Neurol* (2012) 71(1):A5–15. doi:10.1002/ana.23531
38. Stangel M, Penner IK, Kallmann BA, Lukas C, Kieseier BC, Gold R. Multiple sclerosis decision model (MSDM): Entwicklung eines Mehrfaktorenmodells zur Beurteilung des Therapie- und Krankheitsverlaufs bei schubförmiger Multipler Sklerose. *Akt Neurol* (2013) 40(09):486–93. doi:10.1055/s-0033-1349643

**Conflict of Interest Statement:** PT and CR were funded by the Swiss National Science Foundation (PNT: PP00P1_150739, CRSII3_141965, and 00014_165884, CCR: 105314_152891, CRSII3_141965, and 320030_143443). GS is supported by the Heart and Stroke Foundation Distinguished Clinician-Scientist and Mid-Career Awards following an open peer-reviewed competition. JM is an employee of Roche Farma Spain. MS, MT, and XM—no disclosures.

# E. Appendix to Study 5

*Article*

MDM Policy & Practice

*MDM Policy & Practice*
1–12

DOI: 10.1177/2381468319855642
journals.sagepub.com/home/mdm
SAGE

## Traffic Lights Intervention Reduces Therapeutic Inertia: A Randomized Controlled Trial in Multiple Sclerosis Care

**Gustavo Saposnik, Muhammad Mamdani, Xavier Montalban, Maria Terzaghi, Berenice Silva, Maria Laura Saladino, Philippe N. Tobler*, and Fernando Caceres***

**Abstract**

**Background:** Therapeutic inertia (TI) is a common phenomenon among physicians who care for patients with chronic conditions. We evaluated the efficacy of the traffic light system (TLS) educational intervention to reduce TI among neurologists with MS expertise. **Methods:** In this randomised, controlled trial, 90 neurologists who provide care to MS patients were randomly assigned to the TLS intervention ($n = 45$) or to the control group ($n = 45$). The educational intervention employed the TLS, a behavioral strategy that facilitates therapeutic choices by facilitating reflective decisions. The TLS consisted in a short, structured, single session intervention of 5-7 min duration. Participants made therapeutic choices of 10 simulated case-scenarios. The primary outcome was a reduction in TI based on a published TI score (case-scenarios in which a participant showed TI divided by the total number of scenarios where TI was possible ranging from 0 to 8). **Results:** All participants completed the study and were included in the primary analysis. TI was lower in the TLS group (1.47, 95% CI 1.32-1.61) compared to controls (1.93; 95% CI 1.79-2.08). The TLS group had a lower prevalence of TI compared to controls (0.67, 95% CI 0.62-0.71 vs. 0.82, 95% CI 0.78-0.86; $p = 0.001$). The multivariate analysis, adjusted for age, specialty, years of practice, and risk preference showed a 70% reduction in TI for the TLS intervention compared to controls (OR 0.30; 95% CI 0.10-0.89). **Conclusions:** In this randomized trial, the TLS strategy decreases the incidence of TI in MS care irrespective of age, expertise, years for training, and risk preference of participants, which would lead to better patient outcomes.

This paper is published as Saposnik G, Maurino J, Sempere AP, Terzaghi MA, Ruff CC, Mamdani M, Tobler PN, Montalban X. Overcoming Therapeutic Inertia in Multiple Sclerosis Care: A Pilot Randomized Trial Applying the Traffic Light System in Medical Education. Front Neurol.,8:430, 2017.

*Article*

MDMP&P Policy & Practice

*MDM Policy & Practice*
1–12

DOI: 10.1177/2381468319855642
journals.sagepub.com/home/mdm
SAGE

# Traffic Lights Intervention Reduces Therapeutic Inertia: A Randomized Controlled Trial in Multiple Sclerosis Care

**Gustavo Saposnik, Muhammad Mamdani, Xavier Montalban, Maria Terzaghi, Berenice Silva, Maria Laura Saladino, Philippe N. Tobler*, and Fernando Caceres***

**Abstract**

**Background:** Therapeutic inertia (TI) is a common phenomenon among physicians who care for patients with chronic conditions. We evaluated the efficacy of the traffic light system (TLS) educational intervention to reduce TI among neurologists with MS expertise. **Methods:** In this randomised, controlled trial, 90 neurologists who provide care to MS patients were randomly assigned to the TLS intervention ($n = 45$) or to the control group ($n = 45$). The educational intervention employed the TLS, a behavioral strategy that facilitates therapeutic choices by facilitating reflective decisions. The TLS consisted in a short, structured, single session intervention of 5-7 min duration. Participants made therapeutic choices of 10 simulated case-scenarios. The primary outcome was a reduction in TI based on a published TI score (case-scenarios in which a participant showed TI divided by the total number of scenarios where TI was possible ranging from 0 to 8). **Results:** All participants completed the study and were included in the primary analysis. TI was lower in the TLS group (1.47, 95% CI 1.32-1.61) compared to controls (1.93; 95% CI 1.79-2.08). The TLS group had a lower prevalence of TI compared to controls (0.67, 95% CI 0.62-0.71 vs. 0.82, 95% CI 0.78-0.86; $p = 0.001$). The multivariate analysis, adjusted for age, specialty, years of practice, and risk preference showed a 70% reduction in TI for the TLS intervention compared to controls (OR 0.30; 95% CI 0.10-0.89). **Conclusions:** In this randomized trial, the TLS strategy decreases the incidence of TI in MS care irrespective of age, expertise, years for training, and risk preference of participants, which would lead to better patient outcomes.

**Keywords**

decision making, disease-modifying therapy, educational intervention, multiple sclerosis, randomized clinical trial

Date received: February 8, 2019; accepted: May 3, 2019

The management of multiple sclerosis (MS) evolves with the availability of new disease-modifying agents, varying dosage forms (oral, injectable, infusion), and different safety and efficacy profiles. As a result, physicians who care for MS patients face more complex therapeutic decisions when considering individual patients, number of relapses, activity on brain imaging, and the need to escalate therapy. Moreover, many MS patients remain undertreated.[1–3]

Therapeutic inertia (TI) corresponds to the absence of treatment initiation or intensification when treatment goals are unmet. It affects more than 50% of clinicians caring for patients with chronic conditions.[1,4–6] Insufficient knowledge integration and knowledge-to-action gaps are among the most common explanations for suboptimal therapeutic decisions. However, TI is a complex process

**Corresponding Author**
Gustavo Saposnik, Neuroeconomics Program, Department Economics, University of Zurich, Switzerland. (gustavo.saposnik@econ.uzh.ch) Department of Medicine (Neurology), St. Michael's Hospital, University of Toronto, 55 Queen St E, Toronto, Ontario M5C 1R6, Canada; Telephone: (416) 864-5155 (saposnikg@smh.ca).

also related to other characteristics of the providers such as lack of knowledge about appropriate goals, high patient volume, and time constraints. Some physicians fail to integrate the available information (e.g., severity of the condition, risk of progression, imaging findings affecting outcomes) with best practice recommendations for a given risk-scenario.[7–10] Furthermore, physicians have limited training in risk management and formal learning in medical decision making.[11–13] All of these factors may contribute to TI and undertreatment.

Educational interventions have been designed to optimize knowledge integration and bridge knowledge-to-action gaps for complex medical decisions (e.g., diagnostic challenges, varying risk categories, availability of multiple agents with a broad range of safety/efficacy ratio).[11] One such intervention is the traffic lights system (TLS). The TLS emerged as risk categorization and warning strategy to reduce human errors that relies on a relatively automatic, well-established, and cross-cultural concept to increase the chance of an optimal course of action.[14,15] The TLS facilitates the decision-making process using traffic light terminology, which creates a link between a color, representing a risk level, and an action: red light ("high risk"/"stop and think"), yellow light ("intermediate risk"/"reassess soon"), and green light ("low risk"/"continue the same strategy"). For example, studies showed that the TLS facilitates healthier food choices by interrupting automatic behavior and triggering a reevaluation processes.[16] A functional magnetic resonance imaging study showed that TLS labels enhance the coupling between brain regions associated with valuation (i.e., ventromedial prefrontal cortex) and self-control.[17] Evidence from the literature suggests that medical decisions leading to TI are likely related to knowledge-to-action gaps.[7] The design and application of the TLS as an educational intervention provides a unique opportunity to overcome knowledge-to-action gaps in MS care.[18]

In a previous pilot study, we assessed the feasibility and potential efficacy of a TLS educational intervention in 25 neurologists from Spain.[19] TI was present in 72.0% of participants in at least one case scenario. The primary feasibility outcome, the completion rate of the study, was 100% (25/25 participants). While not powered to detect a significant difference between groups, our pilot study demonstrated a nonsignificant reduction in TI for the targeted intervention group relative to the control group (22.6% v. 33.9% postintervention; odds ratio [OR] = 0.57; 95% confidence interval [CI] = 0.26 to 1.22).[19] The TLS also showed a high usability score (74.7; 95% CI = 70.1 to 79.2) when tested in a larger study comprising neurologists from Argentina, Chile, and Canada.[20]

In the present study, we evaluated the efficacy of our simple and pilot-tested educational TLS intervention for reducing TI among neurologists who routinely provide care of MS patients.

## Methods

### *Study Design and Participants*

The overarching goal of the TLS intervention was to facilitate the integration of gaps between risk stratification and treatment decisions (initiation or escalation) in MS care. Specifically, our randomized parallel trial tested the efficacy of an educational TLS intervention (active group) against usual care (control group) for reducing TI in the management of MS (Figure 1). Participants were randomly assigned (1:1 ratio) to the TLS or the control group by Qualtrics (Qualtrics.com; Online Appendix, Figure e1). Allocation concealment was implemented in Qualtrics, so that participants did not know which intervention they were allocated to after completing the initial demographic information. Investigators were also blinded to the treatment allocation. Participants were recruited by automatic e-mail invitation from the study platform. The Institute of Neuroscience Buenos Aires and the Argentinian Neurological Society facilitated the mailing of information to potential participants who met the inclusion criteria.

Division of Neurology, Department of Medicine (GS) and Decision Neuroscience Unit, Li Ka Shing Knowledge Institute (GS, MT), St. Michael's Hospital, University of Toronto, Toronto, Canada; Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Zurich, Switzerland (GS, PNT); Li Ka Shing Centre for Healthcare Analytics Research and Training (LKS-CHART), Toronto, Toronto, Canada (MM); Department of Neurology-Neuroimmunology, Centre d'Esclerosi Múltiple de Catalunya (CEMCAT), Hospital Universitari Vall d'Hebron, Barcelona, Spain (XM); Institute of Neuroscience Buenos Aires (INEBA), Buenos Aires, Argentina (BS, MLS, FC). Prof. Philippe Tobler was funded by the Swiss National Science Foundation (PP00P1_150739 and 100014_165884). Dr. Gustavo Saposnik is a neurologist educated in Argentina, currently practicing neurology in Canada supported by the HSF Scientist Award. Drs. Mamdani, Caceres, Saladino, Silva, Terzaghi, and Montalban have no conflicts to disclose. The study was sponsored by the Institute of Neuroscience Buenos Aires (INEBA) with endorsement by the Sociedad Argentina de Neurologia (SNA) and NeuroEconSolutions, and funded by an operating grant from Roche. Neither the SNA nor Roche were involved in the design, execution, analysis, and interpretation or reporting of the results.

*Drs. Caceres and Tobler contributed equally.

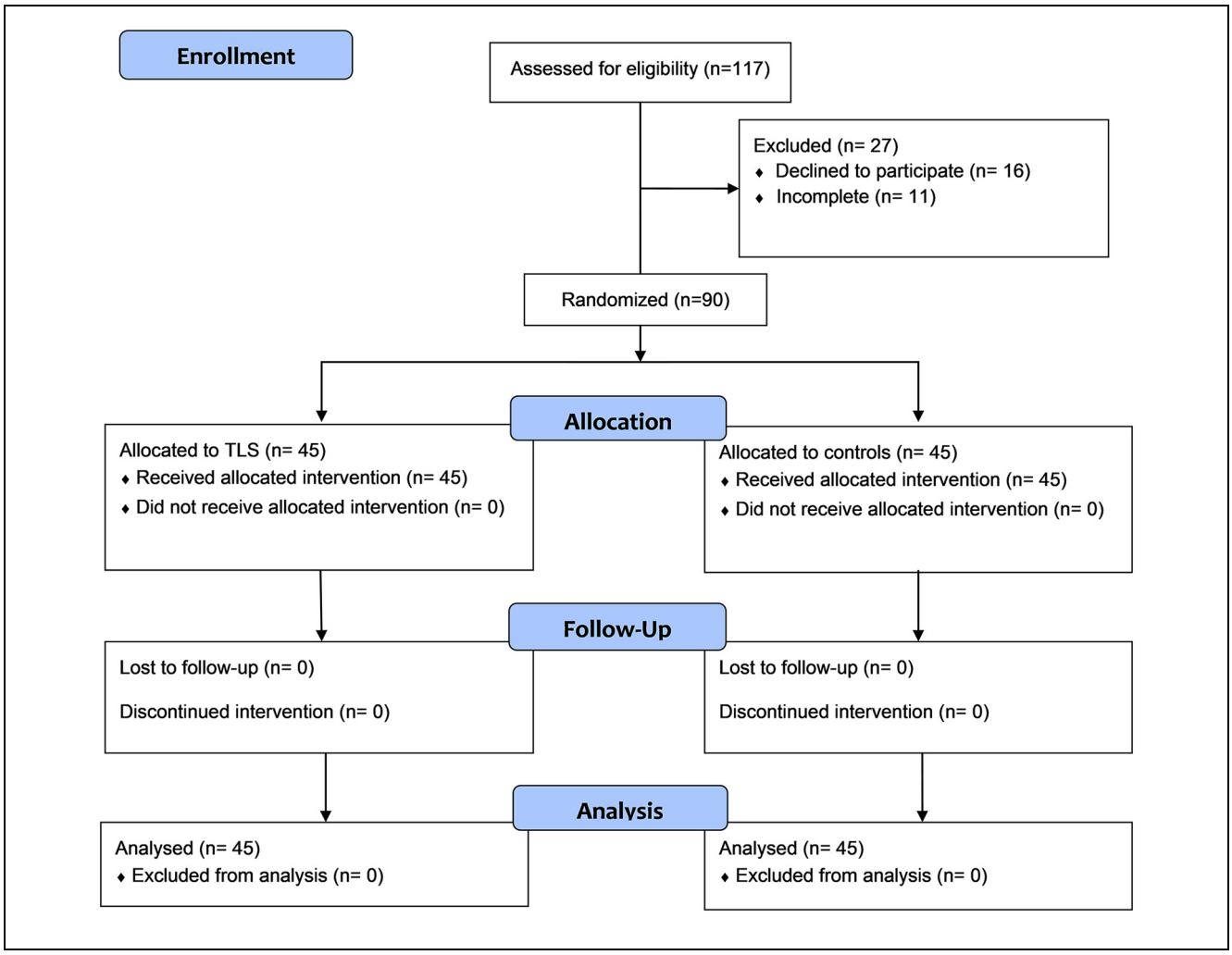



**Figure 1** Consort flow diagram. Of 117 eligible participants, 90 participants were randomized to the educational intervention ($n$ = 45) and control ($n$ = 45). All participants completed the intervention and contributed a complete set of data to the analysis.

## *Data Collection*

Participants answered questions as follows: 1) demographic and practice-based information, 2) behavioral experiments, and 3) 10 case scenarios that assessed TI.

Behavioral experiments used previously established designs to assess participants' risk preferences and tolerance to ambiguity in the health and financial domains.[21,22] In brief, ambiguity aversion is defined as dislike for events with unknown probability compared to events with known probability.[21] For example, in the medical domain, an ambiguity-averse individual would rather choose a treatment where the probability of benefits or side effects is known (even if these are somewhat unfavorable) over one where this probability is unknown (Figure 2A and B). Risk experiments involved determining the subjective value of a risky (50/50) option in terms of a safe (100%) option. By asking participants to indicate the magnitude of the safe option at which they would be indifferent between the two options, we were in a position to determine the point of subjective equivalence of the risky and the safe option. The higher this point, the higher the propensity to take risk. Further details of these experiments were published in previous studies (and appendix).[3]

Participants were also asked to select those MS drugs that they use and then rank them from a list including all available agents approved by the local regulatory body in

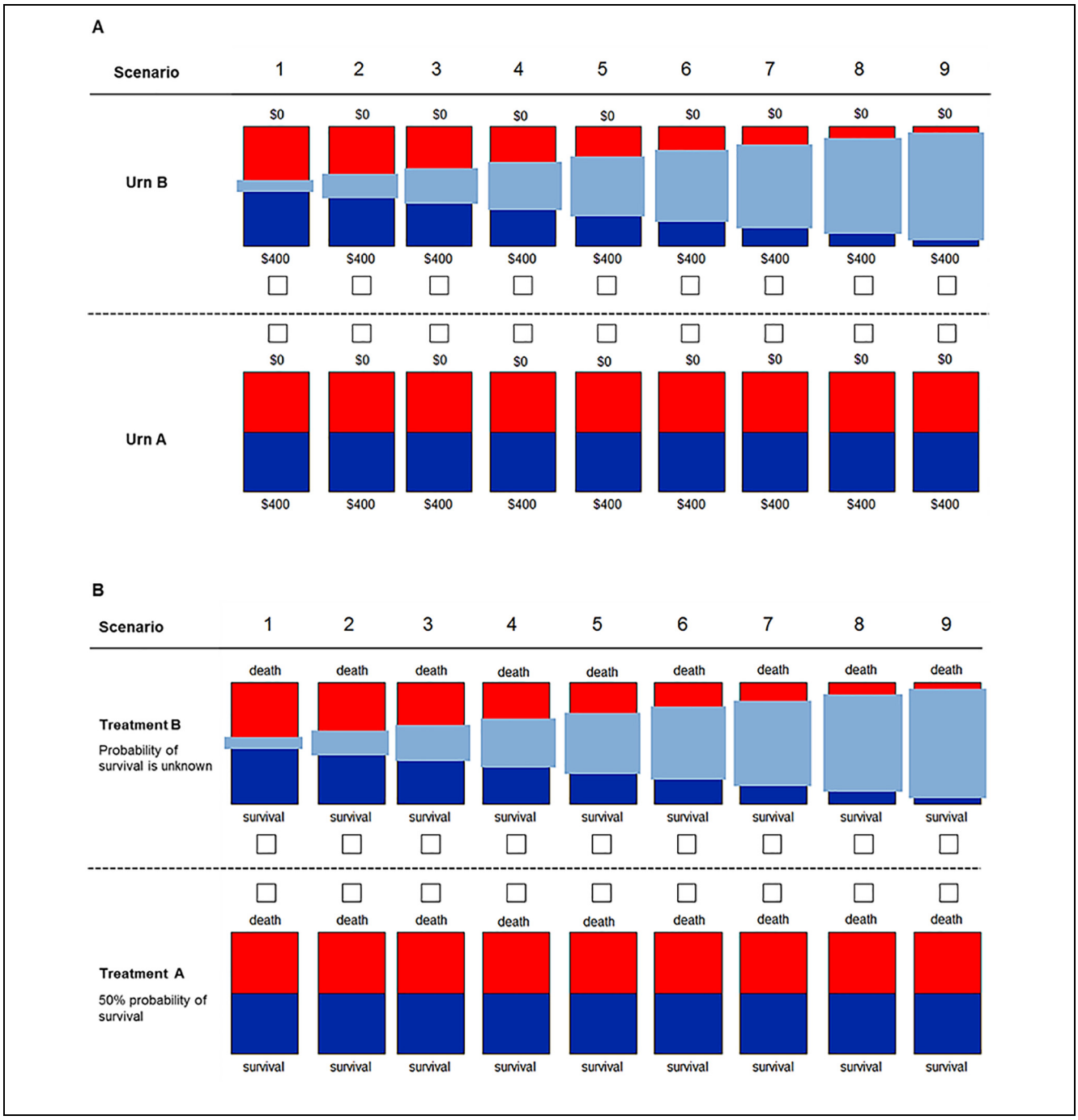



**Figure 2** Experiments to assess ambiguity in the financial and health domains. Participants were told to imagine two different types or urns. For urn type A, they knew that 50% of the balls were red and the other 50% were blue. For urn type B, they did not know the exact proportion of blue to red balls, with the grey bar representing the unknown proportion of balls. For the financial domain (Panel A), participants knew that if they drew a blue ball, they would win the full amount of $400. If they drew a red ball, they would win $ 0. For the health domain (Panel B), participants decided between two treatments for a patient. With "Treatment A," the patient had a 50% probability of survival. With "Treatment B," the exact probability of survival was unknown, with the grey bar representing the unknown probability. In our tasks, participants were asked to choose between one option (presented as two-colored bar) with known 50/50 probability of winning 400 or 0 American dollars (urn A) versus an option with unknown probability of the same outcomes (urn B). Participants who chose the 50/50 options were classified as averse to ambiguity; the remaining participants were classified as tolerant to ambiguity.[1] A similar approach was used to determine aversion to ambiguity in the health domain (Panel B).

Argentina by March 30, 2018. The purpose of this strategy was to compare how TI derived from case scenarios relates with agents used for treatment escalation (e.g., fingolimod, natalizumab, alemtuzumab) in routine clinical practice. At the time of conducting this trial, cladribine and ocrelizumab were not available and therefore not included in case scenarios.

In line with the learning and education literature, vignettes, clinical case scenarios, or "real-world" encounters are regarded as the best simple strategy to evaluate potential cognitive biases and medical decisions.[23] Case scenarios were designed by our research team and MS experts derived from the most common situations in clinical practice as previously reported in our pilot study.[19] Eight case scenarios were designed to assess appropriate treatment initiation or escalation, whereas the remaining two cases were designed to assess overtreatment (defined as treatment escalation when there was low risk of disease progression and no evidence of disease activity).[24–26] Participants from each randomized group were exposed to the same case scenarios. Inclusion criteria comprised neurologists who were actively involved in managing MS patients. Physicians whose practice was primarily in caring for MS patients or who obtained a subspecialty degree were classified as "MS specialists." Physicians who were not practicing neurology or seeing less than one MS patient per month were excluded from the study.

## *Definitions*

For the primary analysis, bad prognosis was defined as the combination of a clinical relapse plus the presence of new brain lesions in follow-up magnetic resonance imaging (MRI) scans or at least one gadolinium-enhancing lesion.[27,28] All high-risk cases included a description of an MRI with more than five new T2 lesions or at least one enhancing lesion.[29] The use of these definitions combining a clinical relapse and MRI activity is consistent with recent evidence regarding the risk of treatment failure among patients receiving interferon-β.[30] Disease progression was defined as at least one point worsening from baseline to 1-year follow-up in the Expanded Disability Status Scale (EDSS) score.[31] Recent meta-analysis confirmed that alemtuzumab, natalizumab, and fingolimod are the best available choices for preventing clinical relapses in patients with relapsing-remitting MS (RRMS).[32] The current treatment option for RRMS include first-line (β-interferons, glatiramer acetate), second-line (fingolimod, cladribine), and third-line (natalizumab, alemtuzumab, ocrelizumab) therapies. For the present analysis, we used the aforementioned scheme according to the current clinical practice.[24–26,33]

## *The Traffic Light System (Figure 3)*

In our study, the TLS was applied to help participants identify scenarios with poor prognosis (high risk of disease progression; Figure 3B). Consequently, participants would be able to identify the "red" traffic light as a warning sign for a high-risk situation, and subsequently escalate treatment. The "yellow" traffic light represents caution for scenarios with either a clinical relapse or some degree of activity on brain imaging (but not both), which requires a reassessment within 6 to 12 months. The control group made therapeutic decisions without being exposed to the TLS intervention as part of the current standard practice. Further details of the TLS intervention are described elsewhere.[19] An example of the presented case scenario is represented in Figure 3C.

## *Outcome Measures*

Therapeutic inertia was the primary outcome of interest, measured as a continuous variable (TI score) and as binary. The TI score was defined as the number of case scenarios in which a participant showed TI (ranging from 0 to 8). The TI score was reported in our previous studies to reflect the magnitude of TI.[19,34] A low TI score represents low TI, whereas a higher score represents higher. A 0.5 point difference in the TI score was deemed as clinically meaningful given the impact in clinical practice. In our study, detecting a difference equal to or greater than 0.50 between groups in the TI score would represent a clinically meaningful improvement. The TI score was derived from case scenarios, which were aligned with the current Argentinian, North American, and European practice recommendations.[24–26]

A reduction in the TI score reflects that participants appropriately switched from a first-line agent (e.g., glatiramer, interferons) to a high-efficacy treatment (e.g., fingolimod as a second-line therapy or monoclonal antibodies as third-line therapies) when clinical and radiological evidence of disease progression.

TI as a binary outcome was defined as the proportion of participants demonstrating TI in at least one of the eight scenarios (prevalence of TI).

## *Statistical Analysis*

We used parametric tests (*t* test and Fisher's exact test) to compare continuous and categorical variables between

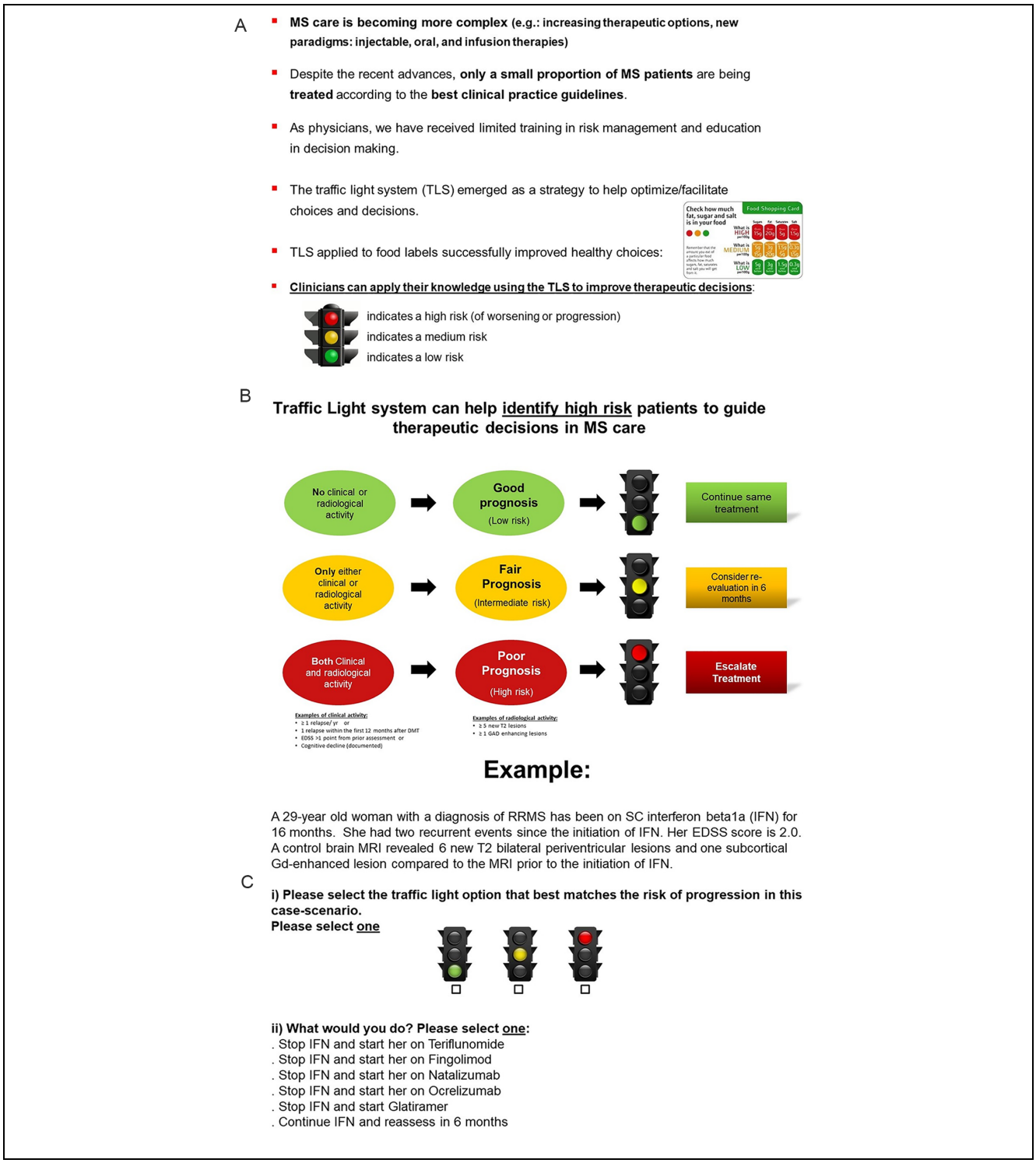



**Figure 3** The TLS intervention. Panel A illustrates background information on the TLS and application to therapeutic decisions. Panel B illustrates how the TLS facilitates the decision-making process using traffic light terminology, which creates a link between a color, representing a risk level, and an action: red light ("high risk"/"stop and think"), yellow light ("intermediate risk"/"reassess soon"), and green light ("low risk"/"continue the same strategy"). Panel C provides a case scenario as an example of those given the participants.

**Table 1** Baseline Characteristics of Participants

| Characteristics | Control ($N = 45$) | Educational Intervention ($N = 45$) | $P$ Value |
|---|---|---|---|
| Age in years (mean ± SD) | 46.8 ± 10.3 | 46.0 ± 10.4 | 0.72 |
| Sex, *n* (%) | | | 0.67 |
| Female | 20 (44.4) | 22 (48.9) | |
| Male | 25 (55.6) | 23 (51.0) | |
| Specialty, *n* (%) | | | 0.82 |
| General neurologists who care for MS patients | 29 (64.4) | 30 (66.7) | |
| MS specialists | 16 (35.6) | 15 (33.3) | |
| Practice setting, *n* (%) | | | 0.37 |
| Public | 17 (37.8) | 13 (28.9) | |
| Private | 28 (62.2) | 32 (71.1) | |
| % time in clinical practice, *n* (%) | | | |
| >75% | 19 (42.2) | 20 (44.4) | 0.83 |
| Years in practice (mean ± SD) | 21.1 ± 10.5 | 19.5 ± 11 | 0.48 |
| MS patients seen per week (mean ± SD) | 21.8 ± 4.5 | 23.0 ± 8.3 | 0.37 |
| Author of a peer-reviewed publication in the last 1 year, *n* (%) | 20 (44.4) | 23 (51.1) | 0.53 |
| Restriction to prescribe MS drugs, *n* (%) | | | 0.83 |
| No restrictions | 28 (62.2) | 29 (64.4) | |
| Treatment escalation, mean drugs (±SD) | 2.91 (1.24) | 3.09 (2.05) | 0.62 |

MS, multiple sclerosis

groups. Linear regression analysis was used to determine the efficacy of the TLS for reducing TI scores in the intervention group versus the control group. Similarly, logistic regression analysis was used to determine the efficacy of the TLS with respect to the proportion of participants with TI in at least one case scenario. We also evaluated the association between restrictions for prescribing DMTs and the number of second- and third-line agents commonly used in clinical practice with the TI score.

For multivariate analysis of individual responses, we constructed multilevel mixed-effects models where participants ($n = 90$) and individual responses ($n = 720$; 90 participants each completing 8 case scenarios) entered as random effects. The aim of this analysis was to evaluate the contribution of individual-specific variables to the variation of TI. Given the findings from our previous studies, we included the following a priori variables: participant age, expertise (MS specialist v. general neurologist), years of experience, and aversion to ambiguity.[3,19] All tests were 2-tailed, and $P$ values $<0.05$ were considered significant. We used STATA 13 (StataCorp LP, College Station, TX) to conduct all analyses.

Further details of the protocol are published in ClinicalTrials.gov (# NCT03134794).

## Results

Out of the 117 neurologists with expertise in MS care who were invited to participate in the study, 90 completed the survey (response rate 76.9%) between April and September 2018. There was representation from all provincial territories. The mean (SD) age was 46.4 (±10.3) years; 48 (53%) were male neurologists. Thirty-one (34.4%) participants primarily focused their practice on MS care. They had 20.3 (±10.9) years of practice and were assessing 22 (±6.6) MS patients per week. Table 1 compares baseline characteristics between groups. Groups did not differ in demographics or in risk preferences as measured by the behavioral risk tasks ($P = 0.40$ for risk preferences and $P = 0.63$ for aversion to ambiguity). There were no differences in treatment escalation at baseline between groups. On average, participants in the TLS group used 3.09 agents for treatment escalation versus 2.91 in the control group ($P = 0.62$). There was no association between participants' restrictions to prescribe MS drugs and TI score ($P = 0.44$) or the prevalence of TI ($P = 0.78$).

Table 2 summarizes the primary and secondary outcome measures at the participant and individual response levels. TI scores were significantly lower in the TLS intervention group than in the control group (1.36, 95% CI = 1.23 to 1.50, v. 2.04, 95% CI = 1.90 to 2.17) after adjustment for the prespecified variables (age, specialty, years of practice, and risk preferences). The observed 0.68 difference between groups in the adjusted TI scores was greater than the minimal clinically meaningful measure of 0.5 to detect an improvement. Similarly, participants in the TLS intervention group had a lower

**Table 2** Multivariate Analysis for the Primary and Secondary Outcome Measures[a]

| Outcome Measures | Control | Intervention | Difference Between Groups | Multivariate Regression Analysis (95% CI) |
|---|---|---|---|---|
| *Primary outcome* | | | | |
| Participant-level analysis | $n = 45$ | $n = 45$ | | |
| TI score, mean (SD) | 1.93 (1.42) | 1.47 (1.42) | (0.46) | −0.68 (−1.24 to −0.11)[b] |
| *Secondary outcome measures* | | | | |
| TI (present v. absent) in at least one case scenario, *n* (%) | 37 (82.2) | 30 (66.7) | (15.5) | 0.30 (0.10 to 0.89)[c] |
| Individual responses | $n = 360$ | $n = 360$ | | |
| TI score, mean (SD) | 1.93 (1.41) | 1.47 (1.41) | (0.46) | −0.68 (−0.87 to −0.48)[d] |
| TI present versus absent, *n* of individual responses/total | 87/360 (24.2) | 66/360 (18.3) | (5.9) | 0.60 (0.41, 0.88)[d] |

CI, confidence interval; OR, odds ratio; TI, therapeutic inertia.
[a]All models adjusted for age, specialty, years of practice, risk preference, and group (intervention v. control).
[b]Derived from linear regression models and expressed in β coefficients (95% CI) with TI score as dependent variable.
[c]Derived from multivariate logistic regression analysis with TI (present v. absent) as dependent variable.
[d]Derived from multilevel mixed effects models expressed as OR (95% CI) for binary outcomes (TI present v. absent) and β coefficients (95% CI) for the TI score.

prevalence of TI compared to controls (63.7%, 95% CI = 58.9% to 68.6%, v. 83.9%, 95% CI = 80.4% to 88.4%) (Figure 4).

The multivariate analysis also revealed a significant reduction in the TI score for the TLS intervention compared to the control group (β = −0.68; 95% CI = −1.24 to −0.11). The multivariate logistic regression analysis showed 70% reduction in the odds of TI for the TLS group (OR = 0.30, 95% CI = 0.10 to 0.89). Specialist status ($P = 0.002$), higher years of experience ($P = 0.007$), and tolerance to ambiguity ($P = 0.043$) were associated with lower TI. The adjusted models showed good discrimination (c-statistic = 0.74) and calibration (goodness-of-fit test $P = 0.52$).

Results were consistent for the analyses of individual responses (Table 2). There were no significant differences between fixed and random effect models (data not shown). Figure 3 represents the relationship between the observed and predicted TI scores (Figure 5A) and stratifies the data by group (Figure 5B), revealing that TI was consistently lower in the intervention group.

Participants who commonly used agents for treatment escalation in their daily practice had lower TI scores (1.48 v. 1.8; $P < 0.01$). Accordingly, participants who do not commonly use agents for treatment escalation in their daily practice benefited from the intervention (for TI score: β = −0.78, 95% CI = −1.03 to −0.52; for TI prevalence: OR = 0.56, 95% CI = 0.36-0.88).

The analysis of individual responses also revealed that for every 100 MS patients with a bad prognosis (e.g., both clinical and radiological evidence of disease activity), there will be over 24 patients who will remain with the same treatment if managed by neurologists without educational intervention (control group). That number would be decreased to 10 patients if treated by neurologists who received the TLS educational intervention.

## Discussion

In the present randomized controlled trial (RCT), we evaluated the efficacy of a newly designed pilot-tested[19] educational intervention to overcome TI among practicing neurologist care for MS patients. We found TI in 7 out of 10 participating neurologists in at least one case scenario. The TLS educational intervention was associated with a 68% reduction in the TI score or 70% reduction in the odds of TI. In other words, participants appropriately choose a higher efficacy treatment (e.g., monoclonal antibodies) instead of continuing with the same agents (e.g., glatiramer, interferon) when clinical and radiological evidence of disease progression. The effect of the educational intervention was similar for all categories of TI scores. Specialist status, years of experience, and tolerance to ambiguity were associated with lower TI. Moreover, selection of common agents used for treatment escalation in participants' routine practice was associated with lower TI scores. More interestingly, the TLS intervention was effective among participants who do not commonly use agents for treatment escalation in their daily practice by showing a significant reduction in TI. Our results were consistent for the

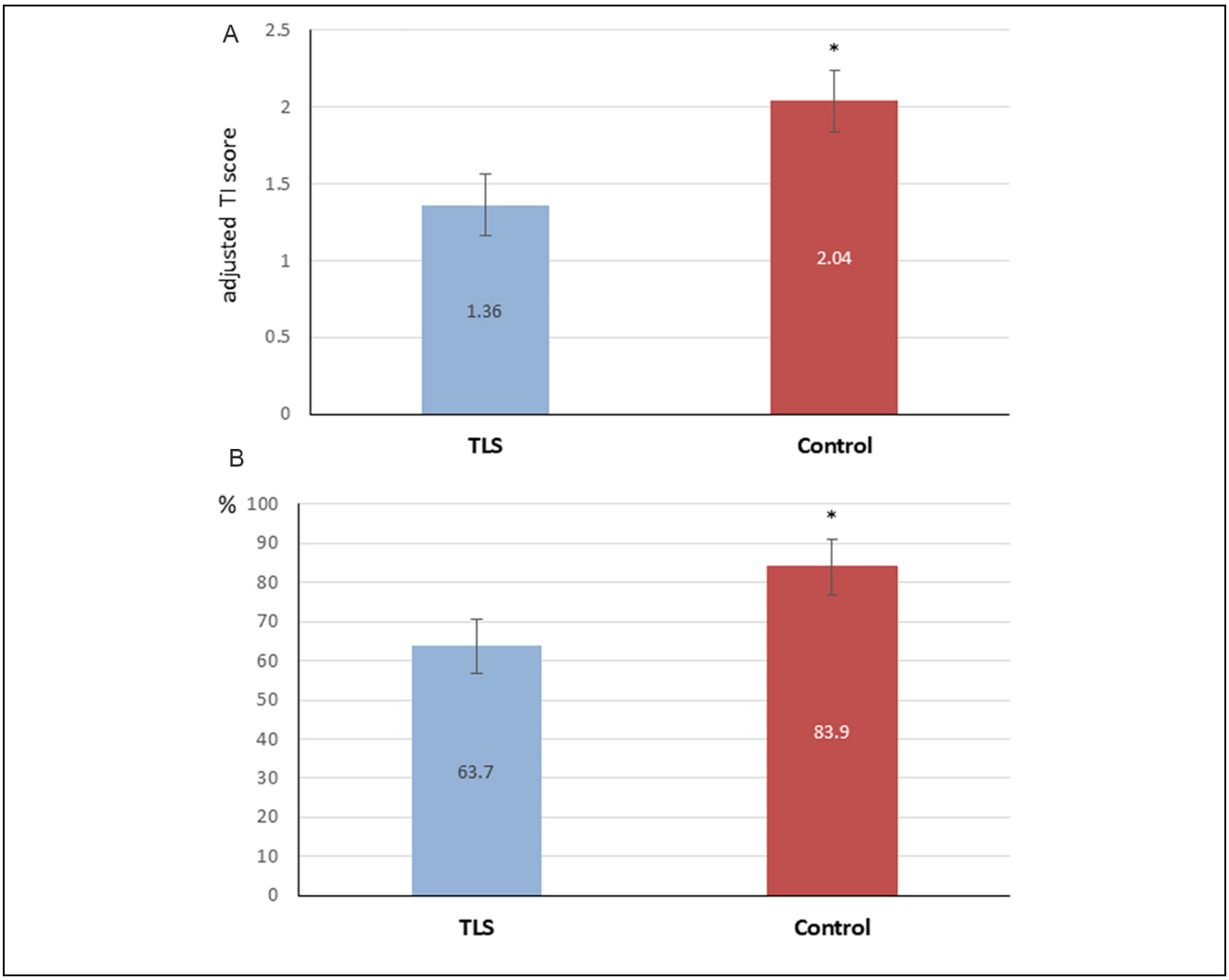



**Figure 4** TLS intervention decreased therapeutic inertia (TI). (A) Comparison of adjusted TI scores in the intervention and control groups. This graph was derived from the multivariate linear regression analysis adjusted for age, years of practice, participants risk preference, and specialty (general neurologist v. MS expert). TI scores were significantly higher in the control group compared to the intervention group (*$P$ value $<0.001$). (B) Comparison of adjusted prevalence of TI between the intervention and control groups. This graph was derived from the multivariate logistic regression analysis adjusted for age, years of practice, participants risk preference, and specialty (general neurologist v. MS expert). The prevalence of TI was significantly higher in the control group compared with the intervention group (*$P$ value $<0.001$).

analysis at the participant level and individual responses for both categories of TI and TI scores.

There are not many proven effective interventions in medical education associated with improvements in clinically meaningful outcomes.[35] A recent systematic review evaluated 302 controlled studies that had investigated the effect of evidence-based educational interventions. Of 85 articles that met the inclusion criteria, 46 (54%) studies were randomized trials, 51 (60%) included postgraduate-level participants. Although the authors evaluated outcomes in multiple domains (e.g., self-efficacy, knowledge, behavior change), none of the studies assessed patients' benefits.[35] In MS care, TI may lead to undertreatment. By extension, the TLS intervention used here may eventually have patient benefits if it reduces TI in MS care.

We used TLS to reduce TI in MS care. Previous authors proposed the TLS to monitor treatment response in patients with RRMS. They included a more

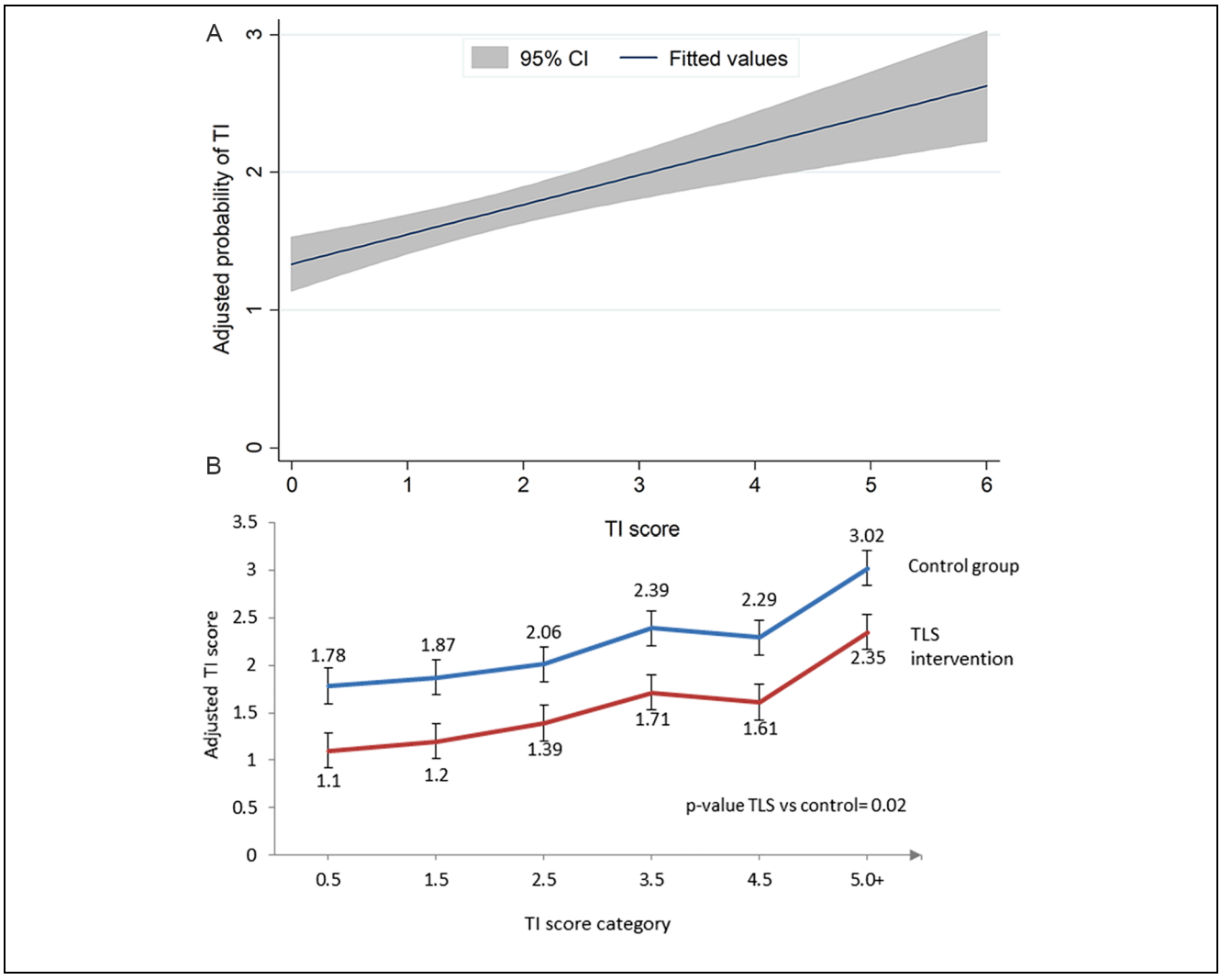



**Figure 5** Adjusted probability of therapeutic inertia (TI). (A) Adjusted probability of TI as a function of TI scores. (B) Adjusted TI score categories stratified by intervention assignment group. The *x*-axis represents categories of the TI to evaluate whether the intervention had a different effect among participants with low, medium, and high TI scores. The *y*-axis represents the TI scores to be able to show the lack of overlap of 95% CI between TLS and controls for each TI category (*P* value TLS v. controls = 0.02). Data derived from multivariate linear regression with TI score as the dependent variable. "I" represent 95% CI error bars.

sophisticated scoring system (0 to 3) for different categories (clinical relapses, evidence of disease progression, cognitive status, and MRI findings) making practical use in daily practice more difficult. This scoring system leads to a decision model that uses the TLS to facilitate therapeutic choices.[36] However, this strategy has not been previously tested in a RCT. The findings of our RCT suggest that the TLS may be a useful medical educational intervention, in keeping with research on the management of obesity, fever in children, blood pressure control, and healthy food choices.[14,15]

Our results have limitations that deserve comment. First, our sample size is relatively small. However, our RCT was designed and powered to detect differences in TI following our pilot study.[19] Second, case scenarios may not necessarily reflect participants' daily practice. It is also possible that general neurologists apply the "first do not harm rule" when not escalating treatment (commonly associated with more severe side effects). Following these arguments, our study underestimates the prevalence of TI in real-life practice given a tightly controlled environment and applicability of treatment

recommendations in our RCT design. Third, we cannot rule out the possibility that unmeasured confounders (e.g., health policy, restrictive prescription rules) may play a role for the studied outcome measures. We controlled for this issue by measuring the prevalence of prescription restrictions in the workplace for each participant. No association was found between prescription restrictions and the outcomes of interest. Fourth, we have no information regarding the sustainability of the TLS effect on TI given the design of our study. Future research will be needed to investigate this question. Finally, the definition of TI applied to MS care may not be widely used. Nevertheless, we used a practical and conservative definition of TI (absence of escalation with the concomitant presence of a clinical relapse and evidence of imaging activity) consistent with our previous studies, which is supported by guidelines showing improvements in clinical outcomes when escalating therapies (i.e., blood pressure and diabetes).[6,24,26]

In the evolving landscape of MS treatment, new and more effective agents with improvements in safety profiles are becoming available.[33] Despite such advancements, many MS patients remained undertreated.[25,26] Several conditions affect the risk of TI, but physicians' factors are regarded as the most influential.[33] Our results revealed that an innovative and highly usable educational intervention may revert the incident risk of TI among neurologists who care for MS patients. This is also supported by the following facts: 1) participants who commonly use agents for treatment escalation had lower TI, 2) the significant reduction in TI for the TLS intervention group among neurologist who do not commonly use agents for treatment escalation, and 3) the relevance of the role of MS specialist in making therapeutic decisions given the lower prevalence of TI.

Our findings have practical clinical and health policy implications, which may not only lead to improving outcomes for patients but also to the implementation of educational interventions in physicians managing high-risk and complex patients. Our intervention has the potential to be translated to other highly prevalent medical conditions, including the management of hypertension, diabetes, and dyslipidemia commonly affecting individuals at high risk of cardiovascular and cerebrovascular diseases.

## Acknowledgments

The authors are most grateful to all physicians participating in the study. We appreciate the support from the Department of Economics at the University of Zurich.

## Supplemental Material

Supplementary material for this article is available on the *Medical Decision Making Policy & Practice* website at https://journals.sagepub.com/home/mpp.

## ORCID iD

Gustavo Saposnik https://orcid.org/0000-0002-5950-9886

## References

1. O'Connor PJ, Sperl-Hillen JAM, Johnson PE, Rush WA, Biltz G. Clinical inertia and outpatient medical errors. In: Henriksen K, Battles JB, Marks ES, Lewin DI, eds. *Advances in Patient Safety: From Research to Implementation (Volume 2: Concepts and Methodology)*. Rockville (MD): Agency for Healthcare Research and Quality; 2005.
2. Ziemssen T, Derfuss T, de Stefano N, et al. Optimizing treatment success in multiple sclerosis. *J Neurol*. 2016;263(6):1053–65. doi:10.1007/s00415-015-7986-y
3. Saposnik G, Sempere AP, Prefasi D, et al. Decision-making in multiple sclerosis: the role of aversion to ambiguity for therapeutic inertia among neurologists (DIScUTIR MS). *Front Neurol*. 2017;8:65. doi:10.3389/fneur.2017.00065
4. Mohan AV, Phillips LS. Clinical inertia and uncertainty in medicine. *JAMA*. 2011;306(4):383.
5. Phillips LS, Branch WT, Cook CB, et al. Clinical inertia. *Ann Intern Med*. 2001;135(9):825–34.
6. Okonofua EC, Simpson KN, Jesri A, Rehman SU, Durkalski VL, Egan BM. Therapeutic inertia is an impediment to achieving the Healthy People 2010 blood pressure control goals. *Hypertension*. 2006;47(3):345–51. doi:10.1161/01.HYP.0000200702
7. Gongora-Ortega J, Segovia-Bernal Y, Valdivia-Martinez Jde J, Galaviz-DeAnda JM, Prado-Aguilar CA. Educational interventions to improve the effectiveness in clinical competence of general practitioners: problem-based versus critical reading-based learning. *BMC Med Educ*. 2012;12:53. doi:10.1186/1472-6920-12-53
8. Kennedy T, Regehr G, Rosenfield J, Roberts SW, Lingard L. Exploring the gap between knowledge and behavior: a qualitative study of clinician action following an educational intervention. *Acad Med*. 2004;79(5):386–93.
9. Djulbegovic B, Beckstead JW, Elqayam S, et al. Evaluation of physicians' cognitive styles. *Med Decis Making*. 2014;34(5):627–37. doi:10.1177/0272989X14525855
10. Kostopoulou O, Sirota M, Round T, Samaranayaka S, Delaney BC. The role of physicians' first impressions in the diagnosis of possible cancers without alarm symptoms. *Med Decis Making*. 2017;37(1):9–16. doi:10.1177/0272989X16644563
11. Dijkstra IS, Pols J, Remmelts P, Brand PL. Preparedness for practice: a systematic cross-specialty evaluation of the alignment between postgraduate medical education and independent practice. *Med Teach*. 2015;37(2):153–61. doi:10.3109/0142159X.2014.929646

12. Monrouxe LV, Grundy L, Mann M, et al. How prepared are UK medical graduates for practice? A rapid review of the literature 2009–2014. *BMJ Open*. 2017;7(1):e013656. doi:10.1136/bmjopen-2016-013656
13. Kostopoulou O, Russo JE, Keenan G, Delaney BC, Douiri A. Information distortion in physicians' diagnostic judgments. *Med Decis Making*. 2012;32:831–9. doi:10.1177/0272989X12447241
14. Sonnenberg L, Gelsomin E, Levy DE, Riis J, Barraclough S, Thorndike AN. A traffic light food labeling intervention increases consumer awareness of health and healthy choices at the point-of-purchase. *Prev Med*. 2013;57(4):253–7. doi:10.1016/j.ypmed.2013.07.001
15. Murphy MS, Baker M. When used appropriately, NICE's traffic light system for febrile children helps clinicians make safe and appropriate decisions. *BMJ*. 2014;348:g2518. doi:10.1136/bmj.g2518
16. Enax L, Krajbich I, Weber B. Salient nutrition labels increase the integration of health attributes in food decision-making. *Judgment Decis Making*. 2016;11(5):460–71.
17. Enax L, Krapp V, Piehl A, Weber B. Effects of social sustainability signaling on neural valuation signals and taste-experience of food products. *Front Behav Neurosci*. 2015;9:247. doi:10.3389/fnbeh.2015.00247
18. Phillips AC, Lewis LK, McEvoy MP, et al. Development and validation of the guideline for reporting evidence-based practice educational interventions and teaching (GREET). *BMC Med Educ*. 2016;16(1):237. doi:10.1186/s12909-016-0759-1
19. Saposnik G, Maurino J, Sempere AP, et al. Overcoming therapeutic inertia in multiple sclerosis care: a pilot randomized trial applying the traffic light system in medical education. *Front Neurol*. 2017;8:430. doi:10.3389/fneur.2017.00430
20. Saposnik G, Tobler PN, Caceres F, et al. Usability of an educational intervention to overcome therapeutic inertia in multiple sclerosis care. *Front Neurol*. 2018;9:522. doi:10.3389/fneur.2018.00522
21. Levy I, Snell J, Nelson AJ, Rustichini A, Glimcher PW. Neural representation of subjective value under risk and ambiguity. *J Neurophysiol*. 2010;103(2):1036–47. doi:10.1152/jn.00853.2009
22. Anderson LR, Mellor JM. Predicting health behaviors with an experimental measure of risk preference. *J Health Econ*. 2008;27(5):1260–74. doi:10.1016/j.jhealeco.2008.05.011
23. Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *JAMA*. 2000;283(13):1715–22.
24. Cristiano E, Alonso R, Pinheiro AA, et al. Argentinean recommendations on the identification of treatment failure in relapsing remitting multiple sclerosis patients. *J Neurol Sci*. 2018;385:217–24. doi:10.1016/j.jns.2018.01.004
25. Montalban X, Gold R, Thompson AJ, et al. ECTRIMS/EAN guideline on the pharmacological treatment of people with multiple sclerosis. *Mult Scler*. 2018;24(2):96–120. doi:10.1177/1352458517751049
26. Rae-Grant A, Day GS, Marrie RA, et al. Practice guideline recommendations summary: disease-modifying therapies for adults with multiple sclerosis: report of the guideline development, dissemination, and implementation Subcommittee of the American Academy of Neurology. *Neurology*. 2018;90(17):777–88. doi:10.1212/WNL.0000000000005347
27. Prosperini L, Mancinelli CR, De Giglio L, De Angelis F, Barletta V, Pozzilli C. Interferon beta failure predicted by EMA criteria or isolated MRI activity in multiple sclerosis. *Mult Scler*. 2014;20(5):566–76. doi:10.1177/1352458513502399
28. Bermel RA, You X, Foulds P, et al. Predictors of long-term outcome in multiple sclerosis patients treated with interferon β. *Ann Neurol*. 2013;73(1):95–103. doi:10.1002/ana.23758
29. Sormani MP, Rio J, Tintore M, et al. Scoring treatment response in patients with relapsing multiple sclerosis. *Mult Scler*. 2013;19(5):605–12. doi:10.1177/1352458512460605
30. Sormani MP, Gasperini C, Romeo M, et al. Assessing response to interferon-β in a multicenter dataset of patients with MS. *Neurology*. 2016;87(2):134–40. doi:10.1212/WNL.0000000000002830
31. Sormani MP, De Stefano N. Defining and scoring response to IFN-β in multiple sclerosis. *Nat Rev Neurol*. 2013;9(9):504–12. doi:10.1038/nrneurol.2013.146
32. Tramacere I, Del Giovane C, Salanti G, D'Amico R, Filippini G. Immunomodulators and immunosuppressants for relapsing-remitting multiple sclerosis: a network meta-analysis. *Cochrane Database Syst Rev*. 2015;(9):CD011381. doi:10.1002/14651858.CD011381.pub2
33. Saposnik G, Montalban X. Therapeutic inertia in the new landscape of multiple sclerosis care. *Front Neurol*. 2018;9:174. doi:10.3389/fneur.2018.00174
34. Saposnik G, Montalban X, Selchen D, et al. Therapeutic inertia in multiple sclerosis care: a study of Canadian neurologists. *Front Neurol*. 2018;9:781. doi:10.3389/fneur.2018.00781
35. Albarqouni L, Hoffmann T, Glasziou P. Evidence-based practice educational intervention studies: a systematic review of what is taught and how it is measured. *BMC Med Educ*. 2018;18(1):177. doi:10.1186/s12909-018-1284-1
36. Stangel M, Penner IK, Kallmann BA, Lukas C, Kieseier BC, Gold R. Multiple Sclerosis Decision Model (MSDM): Entwicklung eines Mehrfaktorenmodells zur Beurteilung des Therapie- und Krankheitsverlaufs bei schubförmiger Multipler Sklerose. *Akt Neurol*. 2013;40(9):486–93. doi:10.1055/s-0033-1349643

# F. Appendix to Study 6





Original Investigation | Medical Education

# Effect of an Education Intervention on Autonomic Response to Clinical Uncertainty in Physicians

## A Randomized Clinical Trial

Gustavo Saposnik, MD[1,2,3,4]

Marcus Grueschow, PhD[2]

Jiwon Oh, MD, PhD[1]

Maria A. Terzaghi, Pharm[1,4]

Pawel Kostyrko, MD[1]

Shruthi Vaidyanathan, MSc[4]

Rosane Nisenbaum, PhD[4,5]

Christian C. Ruff, PhD[2]

Philippe N. Tobler, PhD[2]

[1]Division of Neurology, Department of Medicine, St Michael's Hospital, University of Toronto, Toronto, Ontario, Canada

[2]Zurich Center for Neuroeconomics (ZNE), Department of Economics, University of Zurich, Zurich, Switzerland

This paper has been submitted and accepted for publication in JAMA Netw Open.

# Effect of an education intervention on autonomic response to clinical uncertainty in physicians: A randomized clinical trial

**Running Title:** Effects of an education intervention to overcome clinical uncertainty: A randomized trial

Authors: Gustavo Saposnik, MD [1-4], Marcus Grueschow, PhD[2], Jiwon Oh, MD PhD[1], Maria A. Terzaghi, Pharm[1,4], Pawel Kostyrko, MD[1], Shruthi Vaidyanathan, MSc[4], Rosane Nisenbaum, PhD[4,5], Christian C. Ruff, PhD[2], Philippe N. Tobler, PhD[2]

[1] Division of Neurology, Department of Medicine, St. Michael's Hospital, University of Toronto; Canada
[2] Zurich Center for Neuroeconomics (ZNE), Department of Economics, University of Zurich; Switzerland
[3] Clinical Outcomes and Decision Neuroscience Unit, St Michael's Hospital, University of Toronto, Toronto, Canada
[4] Li Ka Shing Knowledge Institute, St. Michael's Hospital, University of Toronto; Canada
[5] Applied Health Research Centre, Centre for Urban Health Solutions, St Michael's Hospital, Toronto; Canada

Prof. Tobler and Ruff contributed equally

Title character count: 81
Running title character count: 46
Number of references: 50
Number of tables (all black and white): 3
Number of figures (all black and white): 5
No color figures.

Word count abstract: 348
Word count paper: 2993

**Corresponding author:**
Dr. Gustavo Saposnik
PhD Candidate in Neuroeconomics
University of Zurich
Associate Professor, Department of Medicine (Neurology)
St. Michael's Hospital, University of Toronto
55 Queen St E, Toronto, Ontario, M5C 1R6
Tel: (416) 864-5155, Fax: (416) 864-5150
Email: gustavo.saposnik@unityhealth.to

Search terms: multiple sclerosis, disease-modifying therapy, therapeutic inertia, neuroeconomics, decision making, risk

**Key Points:**

**Question:** How physicians handle uncertainty when making live therapeutic decisions?

**Findings:** In this randomized clinical trial 34 neurologists with expertise in multiple sclerosis (MS) were allocated to an educational intervention or usual care to make 680 therapeutic decisions while we simultaneously measure central arousal (acquired by pupillary responses) a proxy for uncertainty. Arousal responsivity was associated with therapeutic inertia (TI). Our educational intervention showed a 31% significant reduction in TI compared to the control group. Our intervention ameliorates TI by reducing arousal responses.

**Meaning:** Our results suggest that arousal response (pupil dilation) is an indicator of how physicians handle uncertainty when making live therapeutic decisions. Pupil dilation also indicates how the educational intervention would alter changes in physician's decisions under uncertainty, thus decreasing TI.

**Abstract:**

**Importance:** Therapeutic inertia (TI) is a common problem in patient care usually triggered by uncertainty. However, limited information is available on how physicians handle therapeutic decisions under uncertainty.

**Objective:** To evaluate the relationship between arousal (pupil dilation), a proxy measure of therapeutic decisions under uncertainty, a proven effective educational intervention and TI.

**Design, Setting, and Participants:** In this randomised controlled trial, we enrolled 34 neurologists with expertise in multiple sclerosis (MS) practicing at 15 outpatient MS clinics in academic and community institutions from across Canada. Participants were randomly assigned to an educational intervention that facilitates treatment decisions (intervention group) or usual care (control group) from December 2017 to March 2018. Participants listened to 20 audio-recorded simulated case-scenarios while pupil responses were assessed by eye-trackers. Arousal was assessed as pupil dilation in time-periods where critical information was provided (T1: clinical data, T2: neurological status, T3: MRI data).

**Interventions:** Traffic-light system based educational intervention (TLS) vs. usual care (not exposed to the educational intervention). The TLS assist participants identify clinical patterns associated with poor prognosis in MS care, thereby facilitating the decision-making process by exploiting existing associations between colors, representing risk levels, and actions (treatment decisions).

**Main outcome and measures:** TI is defined as lack of treatment escalation despite evidence of disease progression. Pupil dilation was used as a measure of central arousal.

**Results:** Of 38 eligible participants, 34 (89.4%) neurologists completed the study. The mean age (SD) was 44.6 (11.6) years; 38.3% were female. TI was present in 50.0% (17/34) of all participants and was associated with greater pupil dilation ($p<0.05$). For every additional standard deviation of pupil dilation, the odds of TI increased by 51% for T1 (95%CI 1.12-2.03), by 31% for T2 (95%CI 1.08-1.59) and by 49% for T3 (95%CI 1.13-1.97). The intervention significantly reduced TI (RR 31.5%; 95%CI 16.1-47.0). Arousal responses mediated 29.0% of the effect of the educational intervention on TI.

**Conclusions and Relevance:** Individual arousal is an indicator of how physicians handle uncertainty when making live therapeutic decisions. Arousal responses regulate the treatment effect of an educational intervention and therapeutic decisions in MS care.

**Introduction:**

Therapeutic decisions requires an individualized balance of the safety and efficacy profiles of different agents with either imperfect or uncertain information about the outcomes of that choice.[1,2] One decision bias occurring in the context of this uncertainty is therapeutic inertia (TI). TI is characterized by suboptimal decisions to not initiate or intensify treatment when treatment goals are unmet.[3-5] TI affects 60% to 90% of physicians caring for patients with chronic conditions (e.g. hypertension, diabetes, multiple sclerosis). Such suboptimal decisions lead to poorer clinical outcomes and higher health care costs.[3-6] Yet, TI can be reduced: In a randomized controlled trial on multiple sclerosis (MS) care, we found a 70% reduction of TI in neurologists, using a short (less than five minutes) and simple (application of the traffic light system) educational intervention.[7] For MS, overcoming TI corresponds to appropriately switching from a first-line agent (e.g. Glatiramer, Interferons) to a high-efficacy treatment (e.g. Fingolimod, Monoclonal antibodies, etc.) given both clinical and radiological evidence of disease progression.[7-11]

Recent studies showed that decisions under uncertainty are associated with central arousal, as measured by pupil dilation.[12,13] In particular, phasic pupil size increases are associated with suboptimal or erroneous decisions involving high uncertainty.[14,15] However, the role of autonomic arousal during therapeutic decisions under uncertainty is unknown.

We hypothesized that our educational intervention would decrease uncertainty about therapeutic choices, as reflected in decreased arousal responses, and thereby lead to an improvement in therapeutic decisions.[12,14,15]

A better understanding of the link between pupil dilation as a marker of an autonomic arousal response and TI may help to develop more effective educational strategies to overcome suboptimal or erroneous decisions.

In this randomized controlled trial (RCT), we aimed to: i) evaluate the relation between arousal responses and TI, ii) investigate how our previously tested and effective educational intervention[7] affects arousal responses and TI, and iii) assess whether arousal responses mediate the association between the educational intervention and TI (i.e., mediation analysis). We used MS care as an ideal model for complex therapeutic

decisions arising in the management of chronic medical conditions, while taking advantage of our previously tested and effective educational intervention to reduce TI in an RCT.[7] Nevertheless, our findings may also apply to other chronic medical conditions such as hypertension, diabetes, and dyslipidemia.

**METHODS:**

**Study design:** We randomly assigned participants to an educational intervention group or a usual care control group. The educational intervention used the traffic light system (TLS) to reduce TI in the management of MS (details below). The control group made therapeutic decisions without being exposed to the TLS intervention, in line with the current standard practice. Randomized group assignment and allocation concealment was controlled by Qualtrics. Participants were not aware to which group they were allocated. Investigators were also blinded to the treatment allocation. The mean (median) time of study completion was 44.9 (39.9) minutes and participants received 450 Canadian dollars (equivalent to USD 350). Further details of the protocol were published in ClinicalTrials.gov # NCT03134794 and elsewhere.[7]

**Inclusion criteria and Participants:** Neurologists actively caring for MS patients from across Canada were invited to participate by e-mail sent from the Canadian Network of MS Clinics and Neuro-sens (Neuro-sens.com). These networks comprise most MS neurologists in Canada. Participants were recruited from December 13, 2017 to March 2, 2018. Physicians whose practice focuses primarily on caring for MS patients were classified as 'MS specialists'. Physicians seeing less than one MS patient per month were excluded from the study. Participants provided written informed consent. The study was approved by the Research Ethics Board of St. Michael's Hospital-University of Toronto.

**Educational intervention: *The traffic light system (TLS)***. We used our previously established TLS-intervention[7] to assist participants identify clinical patterns associated with poor prognosis (high risk of disease progression based on clinical and imaging profile) in simulated case-scenarios. The TLS facilitates the decision-making

process by exploiting existing associations between colors, representing risk levels, and actions.[7,16-18] For example, the red light represents "high risk" and triggers a "stop and think" action, whereas the green light associates "low risk" with a "continue the same strategy" action. Previous studies showed that the TLS can interrupt automatic behavior leading to more optimal decisions.[19] In our MS care model, the TLS facilitates associating the red traffic light, as a warning sign of disease progression, with switching from a low-efficacy agent (e.g., interferon or glatiramer) to a more effective disease modifying treatment (e.g. monoclonal antibodies).[8-10] Conversely, the green traffic light represents a stable patient following a good clinical course (e.g., no relapse and stable activity on brain imaging) and therefore requiring no immediate therapeutic changes.

***Data collection and study flow*** (**Error! Reference source not found.** 1A)***:*** The study progressed as follows: i) Collection of demographic and practice-based information, ii) completion of behavioral experiments, and iii) completion of 20 simulated and standardized case-scenarios (10 before and 10 after the intervention) to assess TI. We used simulated case-scenarios, reflecting common situations in clinical practice, that were previously validated and designed by our research team and MS experts.[7] All simulated case-scenarios were presented auditorily (via headphones connected to the computer) to avoid interference of visual stimulation and automatic eye movements with pupil responses. The mean ±SD duration of case-scenarios was 35.4 ±7.1 seconds (range 27-50). For each simulated case scenario, we identified three time-periods during which critical information was provided (Appendix, Table e1). Time-period T1 provided critical clinical information (present and previous clinical relapses, type of relapse and/or symptoms). T2 informed about the neurological status of the patient (Expanded Disability Status Scale [EDSS]). T3 provided critical brain imaging information (number and nature of new lesions, GAD enhancing lesions). Time-period T0 before the start of case-scenarios served as baseline, while T4 represented the final segment where standardized questions (e.g., "What would you do? Please select one of the following options") were asked in preparation to the treatment options (**Error!**

**Reference source not found.** 1A). Compared to the baseline T0, we expected an arousal response during critical information periods T1-T3 and little responding during T4.

Based on previously reported associations of risk and ambiguity aversion with TI[20,21], we considered also their relation to arousal responses and TI, using established measures in the financial domain.[22,23] Further details of these measures were published previously (see Appendix).[20]

**Outcome measures:** TI is defined as lack of treatment escalation despite evidence of disease progression. Following best-practice guidelines, we defined disease progression as the combination of a clinical relapse plus the presence of five or more new lesions (T2 or Flair sequences) or at least one gadolinium-enhancing lesion in follow-up magnetic resonance imaging (MRI) scans.[8,9] Using the combination of clinical relapse and MRI activity is consistent with evidence regarding the risk of treatment failure in patients receiving interferon-β.[24] For each of our case-scenarios, we determined TI as a binary variable (present vs absent; primary outcome). A secondary outcome included TI greater than or equal to 25% of responses, meaning that participants did not escalate treatment when recommended in at least 1 out of 4 simulated case-scenarios.[7-10]

**Experimental procedures:** The study was conducted in an ambulatory clinic-type setting to increase ecological validity. Room temperature, light conditions (100 lumens), and participants' sitting positions were held constant. Pupil time-series were z-scored within each participant, to allow comparison of pupil dilation between and within simulated case-scenarios, critical time-periods, and participants. The average pupil size (measured at T0, i.e., 1500ms - 500ms before scenario onset) was taken as pupil baseline.[25] For each simulated case-scenario, we estimated arousal responses by subtracting the average baseline pupil diameter from the peak pupil dilation during each critical time-period (T1-T3, see figure 1B).[13,26] Further details are explained in the appendix.

**Sample size:** Given our main goal and the application of our proven effective educational intervention[7] we conducted a post-hoc calculation. The power to determine differences in phasic pupil response in relation to TI was 99%.

**Statistical analysis:** We applied three analytical approaches: i) comparison of arousal responses across critical time-periods, ii) treatment-effect analysis evaluating the association between arousal responses and TI, and iii) a mediation analysis to assess how the association between individual participant characteristics and TI may be mediated by arousal responses. For (i) we used non-parametric tests (Wilcoxon and Mann-Whitney tests for continuous and categorical variables, respectively). For (ii) we compared high vs. low arousal between groups stratified by intervention period (pre- vs. post-intervention). We used generalized estimating equations (GEE) to assess relationships between the variables of interest with TI accounting for clustering (repeated observations on participants). This analysis controlled for the pre-defined explanatory variables age, specialist status (MS expert vs. general neurologists), years of practice, risk preferences, and ambiguity aversion as identified in our previous research.[20] To test the treatment effect of the educational intervention (TLS), we used difference-in-differences models (also called untreated control group design with pre- and post-test).[27] This allowed us to measure the treatment effect of our intervention by comparing the change over time (post-test minus pre-test performance) between the intervention and control group. Pupil dilation for each participants and case-scenario was tested as a mediator (see below). As a result, we were able to evaluate whether the benefits of the educational intervention on TI were mediated by individual arousal responsivity.

Mediation analysis (iii) is a technique commonly used in the social sciences to explain a relation between an independent variable (e.g. demographic variables) and an outcome via a third variable (called 'mediator').[28,29] The greatest value of mediation analysis in RCT data is that it can establish whether the effects of the intervention (or any independent variable preceding the outcome of interest) on the outcome are mediated by the another standardized measured covariate. Here we measured whether the effect of our intervention on TI is mediated by pupil-indexed arousal.

All tests were 2-tailed, and p-values <0.05 were considered significant. We used STATA 13 (College Station, TX: StataCorp LP) and SAS to conduct all analyses.

**RESULTS:**

**Participant characteristics:** Thirty-eight neurologists with expertise in MS were invited to participate in the study. The cooperation rate was 89.5% (n=34/38); the completion rate 100%. Of the 34 participants who completed the study, pupillary data was available in 30 (88.2%) participants. Two participants in each group had incomplete or missing pupil data (CONSORT Figure). The mean age (SD) was 44.6 (± 11.6) years; 13 participants (38.2%) were female. Table 1 summarizes baseline characteristics of the study population. Participants had a mean of 12.5 (± 12) years of experience and assessed 23.1 (± 16) MS patients per week. Participants showed risk-neutrality in our measures of risk attitudes and the minimal safe amounts participants preferred over the 50/50 gamble did not differ between groups [(Mann-Whitney p-value=0.14; control: $ 175.7 (± 45.2); intervention: 211 (± 78.8)]. Nineteen (55.9%) participants showed aversion to ambiguity in the financial domain.

TI was present in 50.0% (17/34) of participants in at least one case-scenario, representing 7.7% (42/544) of all individual responses that assessed TI. There was no difference in baseline pupil data between groups for each simulated case-scenario (2.82 mm vs. 2.96 mm; Mann-Whitney p=0.57).

**1) Arousal responses predict therapeutic inertia**

Overall, pupil size increased for each time-period relative to baseline (F-test for linear regression analysis, all p<0.0001). The results remained robust after adjustment for the pre-specified covariates (p<0.0001).

Importantly, the multivariate analysis adjusted for age, sex, MS expertise, risk preferences, ambiguity aversion, pre-vs. post-intervention period, and intervention group by time-period (Table 2) showed that pupil dilation was positively related to TI for all critical time-periods (T1 to T3). For every additional standard deviation of pupil dilation, the odds of TI increased by 51% for T1 (OR 1.51; 95%CI 1.12-2.03), by 31% for T2 (OR 1.31;

95%CI 1.08-1.59) and by 49% for T3 (OR 1.49; 95%CI 1.13-1.97). As expected, there was no association between pupil dilation and TI for T4 (OR 1.07; 95%CI 0.86-1.34) as there was no critical information provided. Our results were robust to using secondary outcome measures (Table 2). Together, stronger arousal responses at critical time-periods were associated with stronger TI.

**2) Pre vs. Post-intervention and group allocation (intervention vs. control) affect arousal responses**

Pupil size did not differ significantly between intervention and control groups before the intervention but did so after the intervention for T2, T3, and T4 (**Error! Reference source not found.**. Overall, the multivariable analysis showed that in the post-intervention period, participants in the control group had significantly enlarged pupils compared to the intervention group for T2 (p=0.049), T3 (p=0.004), and T4 (p<0.0001). No difference was observed for T1 (p=0.47). The analysis of dichotomized pupil response (maximum-peak minus mean-baseline greater than or equal to 0.1 z-scored difference as a high arousal vs. below 0.1 z-scored difference -low arousal response) showed similar results (Figure 3). After adjusting for the pre-specified covariates, the multivariate analysis revealed similar results (Table 2).

**3) Effect of the educational intervention on TI**

In our previous study, the educational intervention showed a significant reduction in TI.[7] To assess the presence of a treatment effect on individual TI, we use the difference-in-difference analytical strategy (Table 3). We found that participants in the educational intervention group had a significant 31.5% (95%CI 16.1-47.0) reduction in TI compared to the control group (Figure 4). Linear regression analysis adjusted for participant age, gender, expertise, risk preference and pupil dilation showed that for every therapeutic decision, there was a significant TI decrease of 5.0% (-5.0%, 95%CI -0.8%, -9.3%) in the intervention group. The difference-in-difference analysis revealed no evidence for confounding endogenous effects. Together, these data replicate the previous findings that the TLS intervention reduces TI.[7]

**4) Arousal response mediates the relationship between the educational intervention and TI**

We found strong mediation: arousal responses explained 29.0% of the total mediated intervention effect on TI (Figure 5A). Notably, the mediated effect of arousal on TI was greater than the direct effect of the intervention (full mediator). This is also reflected in the larger values of the multiplication of β coefficients for the indirect effect compared to the smaller β coefficient of the direct effect (intervention on TI) (Figure 5B). The direct effect of the educational intervention on TI was -3.5% (95%CI -9.2, 3.5). Other factors (e.g. age, sex, risk preference), had a non-significant or a negligible effect. Further details are in the appendix. The sensitivity analyses of adding or removing covariates (i.e., risk preference, age, sex, academic practice) revealed no significant changes in the β coefficients (<10%) of the direct or indirect effects (data not shown).

**DISCUSSION**

The role of autonomic arousal for therapeutic decision making has been entirely unexplored. In the present study, we addressed this gap in the framework of therapeutic MS care decisions, with a focus on decisions not to escalate treatment when recommended by best practice guidelines (i.e. TI). We analyzed pupil dilation as an established marker of autonomic arousal[12,13] and found that both continuously measured pupil dilation and dichotomized high vs. low pupil responses were associated with TI. For every additional standard deviation of pupil dilation, the odds of TI increased by 30% to 50% depending on the critical clinical information being provided. Even though our study had low statistical power and Canadian neurologists show comparatively little TI [21], we found that participants in the control group would have reduced TI by almost a third if they would have been randomized to the intervention group. Our data suggest that the intervention may ameliorate TI by reducing arousal responses to critical information. Indeed, pupil dilation mediated the effects of the educational intervention on TI (explaining 29% of the total mediated effect).

Uncertainty may drive rapid pupil-linked arousal responses, affecting behavioral choices[12] and the updating of beliefs with presented evidence.[30] Our results are consistent with this notion. Our educational intervention may reduce arousal by reducing uncertainty and thereby facilitating alternative behavioral strategies.

Specifically, the warning function of a red traffic light makes the need for switching to a more effective agent salient[7] and may concurrently boost confidence in the therapeutic decision.

**What is the relevance of our findings for clinical practice?**

TI commonly affects physicians caring for patients with chronic medical conditions such as MS, diabetes, hypertension, and chronic obstructive pulmonary disease.[31-33] TI has been associated with poorer outcomes, including greater disability, diminished quality of life at the individual level and higher hospitalizations and costs at the health-care level.[6,20,34] Limited knowledge integration and knowledge-to-action gaps may result in automatic responses leading to suboptimal therapeutic decisions.[35]

Our study has several limitations. First, pupil size is not a standard measure in clinical practice. To the extent that arousal plays a significant role in therapeutic decisions, our proof-of-concept suggests that there is value in this measure. Second, arousal responses may have different triggers and effects than the ones we tested. Third, our sample size was small, reducing the precision of our results. Fourth, simulated case-scenarios may not truly reflect therapeutic decisions in clinical practice. Fifth, the traffic light system we used is just one example of an intervention suitable for MS care. Other interventions may be needed and more effective for the management of other prevalent acute and chronic conditions.

Despite these limitations, our conclusions are strengthened by a randomized intervention design showing an arousal-modulated link between an effective therapeutic intervention and therapeutic decisions made by physicians who care for MS patients (Figure 5C).

In conclusion, this study increases our understanding on how physicians make therapeutic decisions under uncertainty. Critical information (i.e., clinical course, neurological status and brain imaging) increases arousal and stronger arousal responses are associated with suboptimal therapeutic decisions (i.e., therapeutic inertia). Moreover, the inertia-reducing effects of an educational intervention appear to be mediated by reduced arousal responses. We are now able to identify physicians making more than 25% suboptimal decisions and to estimate the inertia-reducing benefits of novel educational interventions using the autonomic arousal

response as a marker of their effectiveness. Crucially, this marker is unaffected by demand effects or cognitive biases. Consequently, our findings open avenues to tailor educational interventions and formal risk-assessment training to decision makers (medical students, family doctors, and specialists). In the future, this approach may help optimize treatment decisions for other more prevalent chronic diseases (e.g. hypertension, diabetes) leading to improved medical education and better patients' outcomes.

**Acknowledgements**

The authors are most grateful to all physicians participating in the study.

We are grateful for Prof Kevin Thorpe suggestions regarding the statistical analysis.

We appreciate the support from the Department of Economics at the University of Zurich by facilitating access to the Qualtrics platform.

**Authors contributions:**

GS, CR, PT: study concept and design

GS, PK, MT, SV, RN: acquisition or analysis of data

GS, MG, PT, CR: drafting of the manuscript and figures

**Potential Conflicts of Interest**

Dr. Gustavo Saposnik is supported by the Heart and Stroke Foundation Career Scientist Awards following an open peer-reviewed competition. Dr. Saposnik has received personal compensation for consulting or CME activities from Hoffman LaRoche, Celgene, and Servier Canada.

Maria Terzaghi, Pawel Kostyrko, Shruthi Vaidyanathan, and Rosane Nisenbaum – no disclosures

Dr. Oh has received research funding from the MS Society of Canada, the National MS Society, Brain Canada, Sanofi-Genzyme, Roche, and Biogen. Dr. Oh has received personal compensation for consulting or speaking from EMD-Serono, Genzyme, Biogen, Novartis, Celgene, and Roche.

Prof. Christian Ruff received research funding from the Swiss National Science Foundation (100019L_173248)

Prof. Philippe Tobler received research funding from the Swiss National Science Foundation (PP00P1_150739 and 100014_165884) and Pfizer.

**Study funding:** The study was sponsored by the Division of Neurology at the University of Toronto with the collaboration of Neuro-sens and funded through an unrestricted grant from Roche Canada. The sponsoring, collaborating, or funding entity were not involved in the design, execution, analysis, interpretation of results, or decision to publish the study. Ocrelizumab (manufactured by Roche) was not included as a therapeutic option in any of the simulated case-scenarios.

## References


1. Blumenthal-Barby JS, Krieger H. Cognitive biases and heuristics in medical decision making: a critical review using a systematic search strategy. *Med Decis Making.* 2015;35(4):539-557.
2. Bruck W, Gold R, Lund BT, et al. Therapeutic decisions in multiple sclerosis: moving beyond efficacy. *JAMA Neurol.* 2013;70(10):1315-1324.
3. O'Connor PJ, Sperl-Hillen JAM, Johnson PE, Rush WA, Biltz G. Clinical Inertia and Outpatient Medical Errors. In: Henriksen K, Battles JB, Marks ES, Lewin DI, eds. *Advances in Patient Safety: From Research to Implementation (Volume 2: Concepts and Methodology)*. Rockville (MD)2005.
4. Mohan AV, Phillips LS. Clinical inertia and uncertainty in medicine. *JAMA.* 2011;306(4):383; author reply 383-384.
5. Okonofua EC, Simpson KN, Jesri A, Rehman SU, Durkalski VL, Egan BM. Therapeutic inertia is an impediment to achieving the Healthy People 2010 blood pressure control goals. *Hypertension.* 2006;47(3):345-351.
6. Burks J, Marshall TS, Ye X. Adherence to disease-modifying therapies and its impact on relapse, health resource utilization, and costs among patients with multiple sclerosis. *ClinicoEconomics and outcomes research : CEOR.* 2017;9:251-260.
7. Saposnik G, Mamdani M, Montalban X, et al. Traffic Lights Intervention Reduces Therapeutic Inertia: A Randomized Controlled Trial in Multiple Sclerosis Care. *MDM Policy Pract.* 2019;4(1):2381468319855642.
8. Freedman MS, Selchen D, Prat A, Giacomini PS. Managing Multiple Sclerosis: Treatment Initiation, Modification, and Sequencing. *Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques.* 2018:1-15.
9. Rae-Grant A, Day GS, Marrie RA, et al. Practice guideline recommendations summary: Disease-modifying therapies for adults with multiple sclerosis: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. *Neurology.* 2018;90(17):777-788.
10. Ontaneda D, Cohn S, Fox R. Risk stratification and mitigation multiple sclerosis. *Mult Scler Relat Disord.* 2014;3(5):639-649.
11. Ontaneda D, Tallantyre E, Kalincik T, Planchon SM, Evangelou N. Early highly effective versus escalation treatment approaches in relapsing multiple sclerosis. *Lancet Neurol.* 2019;18(10):973-980.
12. Urai AE, Braun A, Donner TH. Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. *Nat Commun.* 2017;8:14637.
13. Mathot S. Pupillometry: Psychology, Physiology, and Function. *J Cogn.* 2018;1(1):16.
14. Maier ME, Ernst B, Steinhauser M. Error-related pupil dilation is sensitive to the evaluation of different error types. *Biol Psychol.* 2019;141:25-34.
15. Wessel JR, Dolan KA, Hollingworth A. A blunted phasic autonomic response to errors indexes age-related deficits in error awareness. *Neurobiol Aging.* 2018;71:13-20.
16. Enax L, Hu Y, Trautner P, Weber B. Nutrition labels influence value computation of food products in the ventromedial prefrontal cortex. *Obesity (Silver Spring).* 2015;23(4):786-792.
17. Murphy MS, Baker M. When used appropriately, NICE's traffic light system for febrile children helps clinicians make safe and appropriate decisions. *BMJ.* 2014;348:g2518.
18. Zhang X, Liu Y, Gu Y, Wang S, Chen H. Red for "Stop": "Traffic-Light" Nutrition Labels Decrease Unhealthy Food Choices by Increasing Activity and Connectivity in the Frontal Lobe. *Nutrients.* 2020;12(1).
19. Enax L, Krajbich I, Weber B. Salient nutrition labels increase the integration of health attributes in food decision-making. *Judgment and Decision Making.* 2016;11(5):460-471.
20. Saposnik G, Sempere AP, Prefasi D, et al. Decision-making in Multiple Sclerosis: The Role of Aversion to Ambiguity for Therapeutic Inertia among Neurologists (DIScUTIR MS). *Front Neurol.* 2017;8:65.

21. Almusalam N, Oh J, Terzaghi M, et al. Comparison of Physician Therapeutic Inertia for Management of Patients With Multiple Sclerosis in Canada, Argentina, Chile, and Spain. *JAMA Netw Open.* 2019;2(7):e197093.
22. Levy I, Snell J, Nelson AJ, Rustichini A, Glimcher PW. Neural representation of subjective value under risk and ambiguity. *Journal of neurophysiology.* 2010;103(2):1036-1047.
23. Anderson LR, Mellor JM. Predicting health behaviors with an experimental measure of risk preference. *J Health Econ.* 2008;27(5):1260-1274.
24. Sormani MP, Gasperini C, Romeo M, et al. Assessing response to interferon-beta in a multicenter dataset of patients with MS. *Neurology.* 2016;87(2):134-140.
25. Murphy PR, Vandekerckhove J, Nieuwenhuis S. Pupil-linked arousal determines variability in perceptual decision making. *PLoS Comput Biol.* 2014;10(9):e1003854.
26. Privitera CM, Renninger LW, Carney T, Klein S, Aguilar M. Pupil dilation during visual target detection. *Journal of Vision.* 2010;10(10):3-3.
27. Antonakis J, Bendahan S, Jacquart P, Lalive R. On making causal claims: A review and recommendations. *The Leadership Quarterly.* 2010;21(6):1086-1120.
28. MacKinnon DP, Fairchild AJ, Fritz MS. Mediation Analysis. *Annual Review of Psychology.* 2007;58(1):593-614.
29. VanderWeele TJ. Mediation Analysis: A Practitioner's Guide. *Annual Review of Public Health.* 2016;37(1):17-32.
30. de Gee JW, Knapen T, Donner TH. Decision-related pupil dilation reflects upcoming choice and individual bias. *Proc Natl Acad Sci U S A.* 2014;111(5):E618-625.
31. Khunti K, Gomes MB, Pocock S, et al. Therapeutic inertia in the treatment of hyperglycaemia in patients with type 2 diabetes: A systematic review. *Diabetes Obes Metab.* 2017.
32. Ogura M, Harada-Shiba M. Clinical Inertia in the Management of Hypercholesterolemia: What Clinicians Need to do. *J Atheroscler Thromb.* 2016;23(5):552-553.
33. Cooke CE, Sidel M, Belletti DA, Fuhlbrigge AL. Review: clinical inertia in the management of chronic obstructive pulmonary disease. *COPD.* 2012;9(1):73-80.
34. Saposnik G, Sempere AP, Raptis R, Prefasi D, Selchen D, Maurino J. Decision making under uncertainty, therapeutic inertia, and physicians' risk preferences in the management of multiple sclerosis (DIScUTIR MS). *BMC Neurol.* 2016;16:58.
35. Lang ES, Wyer PC, Haynes RB. Knowledge Translation: Closing the Evidence-to-Practice Gap. *Annals of Emergency Medicine.* 2007;49(3):355-363.

**Figure legends**

**Figure 1. Design and basic data**

**A. Study flow.** Participants first answered demographic and practice-based questions and provided risk preferences and ambiguity preferences. Next, they listened to simulated case-scenarios. Each scenario was followed by six therapeutic choices, which remained on the screen until the participant selected one of them. After the first ten simulated case scenarios (pre-intervention), participants were randomized to the intervention (traffic light system) or the control group. Finally, all participants performed another ten simulated case scenarios. The dots after therapeutic choices #1 and #11 represent case scenarios #2-10 and #12-20.

**B. Time- and group-resolved pupil dilation**. Pupil dilation peaked at relevant time-periods both for participants in the control and the intervention group. The black dots represent the peak pupil size within each time-period used to compute pupil responses (pupil peak for each time-period minus mean baseline at time-period 0 (T0). Peaks were determined similarly for both groups across all time-periods and case-scenarios.

**Figure 2. Effects of intervention and group allocation on pupil responses**

Pupil-linked arousal responses (peak minus mean baseline) are shown separately for intervention and control groups, stratified by the intervention period. * $p<0.01$ for the comparison of pupil responses between control and intervention groups. Note that lower responses in the intervention group extend to T4, where no critical information was provided, which may suggest that the protective effect of the intervention extends into the period when participants made decisions under uncertainty.

**Figure 3. Differences in TI by pupil dilation (dichotomized) and intervention groups**

Plot of TI as a function of dichotomized pupil responses (peak minus mean baseline; high responses: ≥0.1 z-scored difference; low responses: <0.1 z-scored difference). The graph was derived from adjusted analysis using GEE accounting for clustering within participants. Dispersion bars indicate 95% confidence intervals.* $p<0.01$ for the comparison of TI between groups. The control group with high arousal response served as reference category. Participants with low pupil responses showed less TI. There was no difference in TI between the intervention and control group for low response individuals. Conversely, high responders showed a significant difference in TI between the intervention and control group.

**Figure 4. Treatment effects using difference-in-differences**

Unadjusted TI before and after intervention allocation group, controlling for confounding endogenous factors. The dotted line illustrates the expected reduction of TI (31.5%, 95%CI 16.1-47.0) if the control group would have received the intervention using difference-in-difference analysis.

**Figure 5. Mediation Analysis**

**A. Schematic representation of generic mediation analysis.** The mediation analysis characterizes the relationship between the independent variable (X) and variables that are related with an outcome of interest (Y). A mediating variable (M) is hypothesized to be intermediate in the relation between an independent variable (X) and an outcome (Y). **B. Structural equation model used for mediation analysis.** The mediation model characterizes the relationship between the independent variable X (Educational intervention) and variables (age, years of practice, MS expert, risk preference) that are related with an outcome of interest Y (TI). The mediating variable M (pupil enlargement from baseline) is hypothesized to be intermediate in the relation between the educational intervention and TI. This figure illustrates the direct, indirect, and total effect of the educational intervention on TI with pupil dilation as the mediator. Pupil responses mediated the effect of the intervention on TI. The Sobel test confirmed that the mediation effect was significant (p=0.029). The effect of covariates on pupil dilation are omitted in the graph for simplicity. **C. Proposed pathways associated with therapeutic inertia in MS care.** Summary schematic illustrating integration of autonomic, behavioral (decision-making), and participant characteristics associated with TI. The reduction in TI (therapeutic decision) by the educational intervention is mediated by pupil dilation. Other covariates may also directly or indirectly influence TI, but in the present study had negligible or non-existent effects (mediation effect <3%).

**Table 1. Baseline characteristics of participants**

| Characteristics | Total (%) n=34 | Control (n=14) | Intervention (n=20) |
|---|---|---|---|
| **Age**, mean (SD), in years | 44.6 (11.6) | 40.5 (8.5) | 47.5 (13.5) |
| **Sex** | | | |
| Female | 13 (38.2) | 6 (42.9) | 7 (35.0) |
| **Practice characteristics** | | | |
| MS specialists | 20 (58.8) | 6 (42.9) | 14 (70.0) |
| General Neurologists who care for MS patients | 14 (41.2) | 8 (57.1) | 6 (30.0) |
| Practice setting: Academic Hospitals | 28 (82.4) | 6 (42.9) | 14 (70.0) |
| Years in practice, mean (SD) | 12.5 (11.8) | 9.4 (9.5) | 14.7 (12.9) |
| MS patients seen per week (≥20) | 15 (44.1) | 4 (28.6) | 11 (55.0) |
| **Author of a peer-reviewed publication in the last 12 months** | 22 (64.7) | 10 (71.4) | 12 (60.0) |
| **Risk preference**, minimal safe amount, mean (SD) in $ | 196.5 (68.5) | 175.7 (45.2) | 211 (78.8) |
| **Ambiguity aversion** | 19 (55.9) | 8 (57.1) | 11 (55.0) |
| **Pupil data** | | | |
| **Pupil size**, mean baseline in mm (SD) | 2.90 (0.87) | 2.82 (0.35) | 2.96 (0.99) |
| **Pupil size**, mean peak in mm (SD) * | 3.27 (1.10) | 3.15 (0.5) | 3.35 (1.35) |
| **Pupil size**, response (peak minus mean baseline), mean in mm (SD)* | 1.60 (1.42) | 1.69 (1.34) | 1.54 (1.47) |

Numbers in brackets indicate percentages, unless otherwise indicated.
Pupil data reflect averages across the study after interpolation and z-scored.
* only including critical time-periods T1, T2, T3
There were no differences in baseline characteristics between groups

**Table 2. Relationship between pupil dilation by critical time-periods and therapeutic inertia**

| **Outcome: Therapeutic Inertia** | **Model for T1**<br>OR (95%CI) | **Model for T2**<br>OR (95%CI) | **Model for T3**<br>OR (95%CI) | **Model for T4**<br>OR (95%CI) |
|---|---|---|---|---|
| Significance of time-period | Clinical presentation | Functional status | MRI findings | Standardized question |
| Maximum pupil dilation minus baseline for TI indicator | 1.51 (1.12-2.03) | 1.31 (1.08-1.59) | 1.49 (1.13-1.97) | 1.07 (0.86-1.34) |
| Maximum pupil dilation minus baseline for TI >25% | 1.53 (1.11-2.12) | 1.33 (1.07-1.63) | 1.51 (1.13-2.00) | 1.08 (0.86-1.36) |
| Maximum pupil dilation minus baseline (aggregated results T1 to T3) for TI indicator | 1.47 (1.24-1.74) | | | NA |
| Maximum pupil dilation minus baseline (aggregated results T1 to T3) for TI >25% | 1.49 (1.19-1.87) | | | NA |

This table represents the results of the multivariate models adjusted for age, sex, MS expertise, risk preferences, ambiguity aversion, pre-vs. post-intervention period, and intervention group. Across all case scenarios, pupil responses (maximum dilation minus mean baseline) were associated with TI for all critical time-periods (T1 to T3). For every additional standard deviation of pupil dilation, the odds of TI increase by 31% to 51%. As expected, there was no association between pupil dilation and TI for time-period 4 (standardized questions at the end of the presentation for each case-scenario). The analysis of aggregated pupil data for T1, T2 and T3 revealed the for every additional standard deviation of pupil dilation, the odds of TI increase by 47% to 49% for the TI indicator or for TI greater than or equal to 25% of responses, respectively.

**Table 3. Effect of group on pupil dilation: Difference-in-difference analysis**

| **Outcome: Pupil response** | **Model for T1** | **Model for T2** | **Model for T3** | **Model for T4** |
|---|---|---|---|---|
| Time-period | Clinical presentation | Functional status | MRI findings | Standardized question |
| **Pupil response (peak minus mean baseline size)** | | | | |
| | β coef. (95%CI) | β coef. (95%CI) | β coef. (95%CI) | β coef. (95%CI) |
| Intervention vs. control (unadjusted) | - 0.23 (-0.64, 0.19) | -0.46 (-0.91, -0.02) | -0.63 (-1.04, -0.22) | -0.75 (-1.12, -0.38) |
| Intervention vs. control (adjusted) | - 0.23 (-0.64, 0.19) | -0.46 (-0.90, -0.01) | -0.63 (-1.04, -0.22) | -0.75 (-1.12, -0.37) |
| Intervention vs. control (adjusted) | - 0.68 (-1.10, -0.25) | | | NA |

This table represents the results of the univariate and multivariate models adjusted for age, sex, MS expertise, risk preferences, ambiguity aversion, pre-vs. post-intervention period, and intervention group. Across all case scenarios, participants in the intervention group showed significantly smaller pupil responses (difference in difference post-period minus pre-period) compared to controls for time-periods T2, T3 and T4. For example, pupil responses were reduced by an estimated 0.46 mm for T2, 0.63 mm for T3 and 0.75 mm for T4 in the intervention group compared to the control group. This is equivalent to a significant relative reduction of pupil responses of 29.3% for T2, 32.5% for T3 and 38.7% for T4 in the intervention group compared to the control group.

**Figure 1A: Study flow**

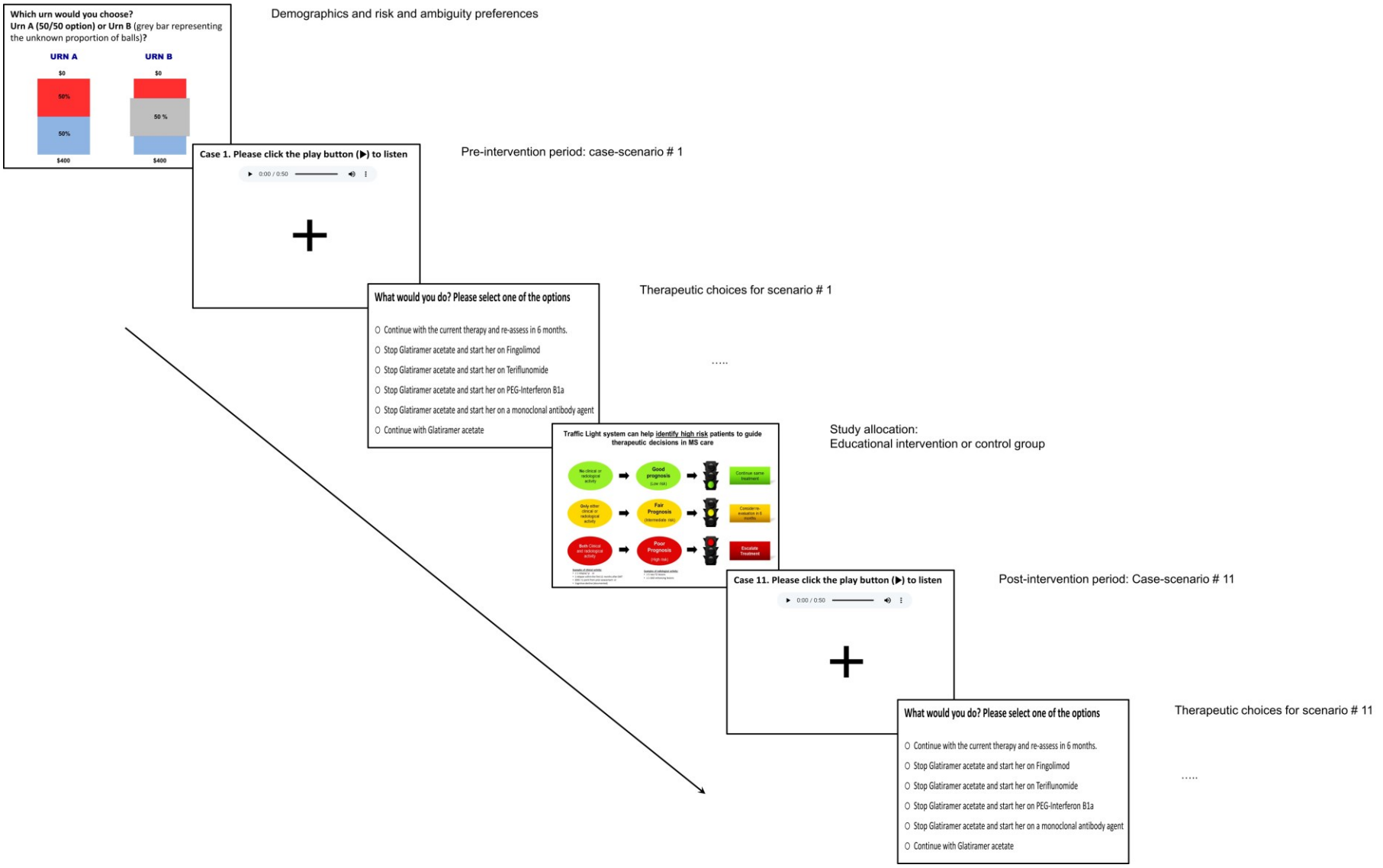


**Figure 1B: Time- and group-resolved pupil dilation**

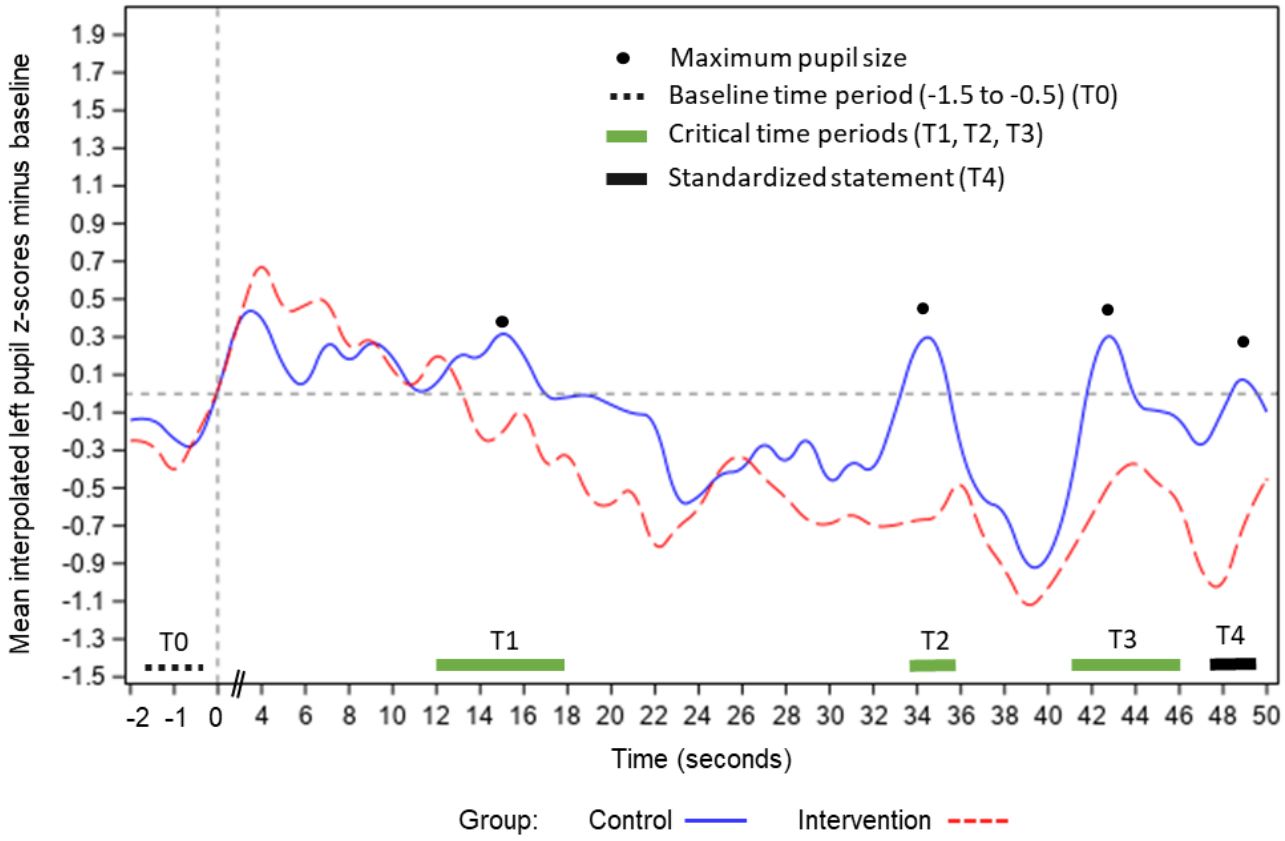


**Figure 2: Effects of intervention and group on pupil responses**

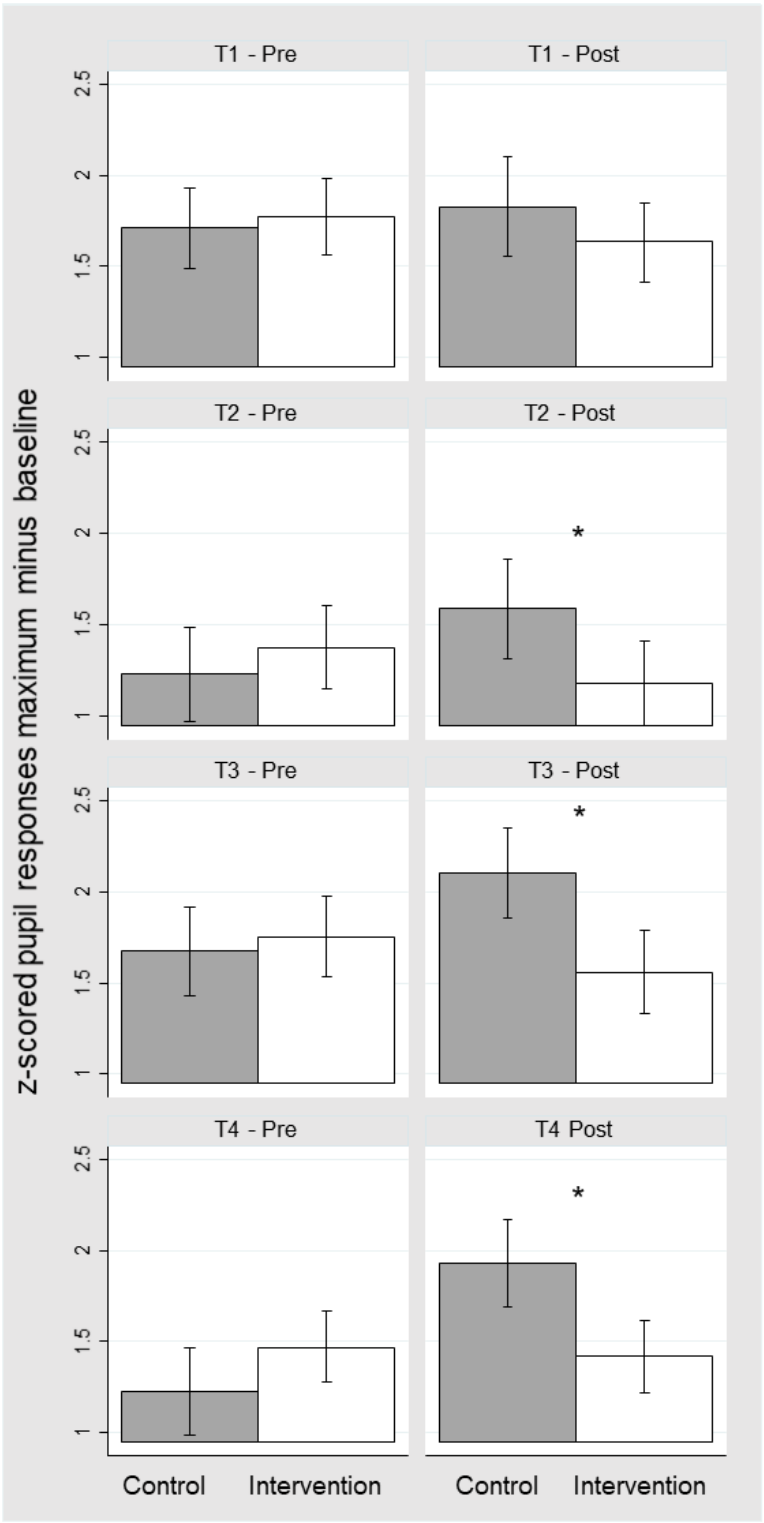
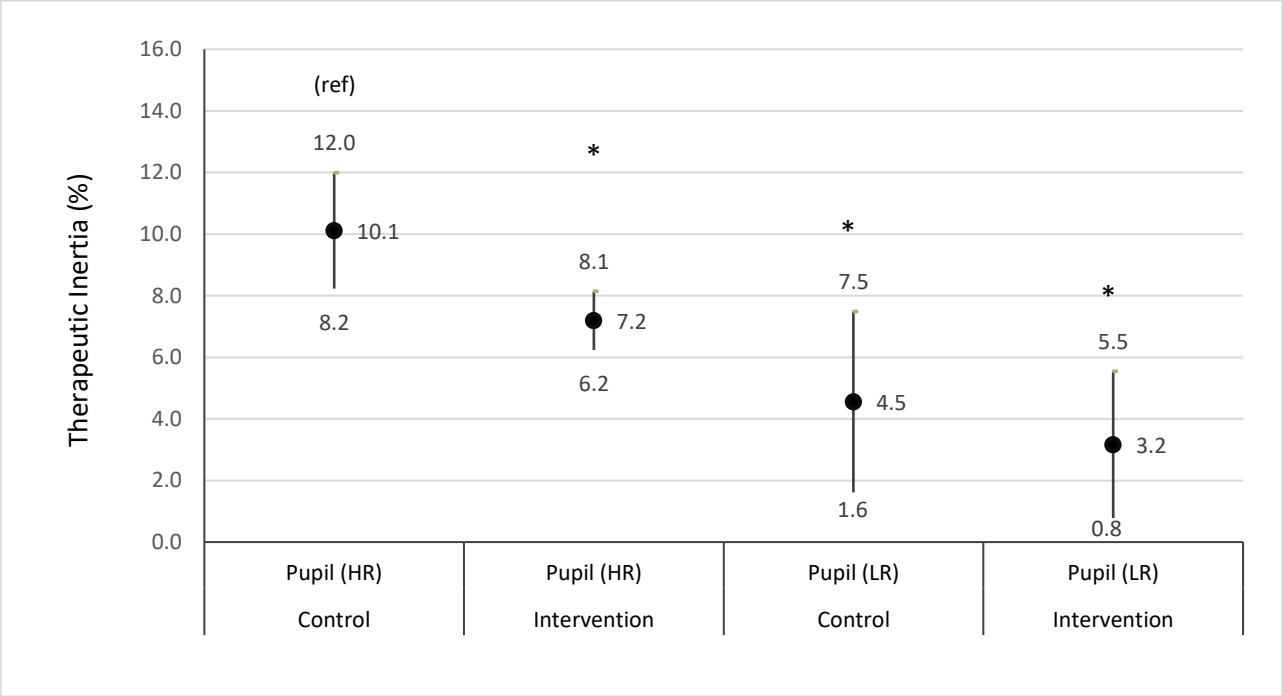


**Figure 3: Differences in TI by pupil dilation (dichotomized) and intervention groups**




HR: high pupil response; LR: low pupil response

**Figure 4: Treatment effects using difference-in-differences**

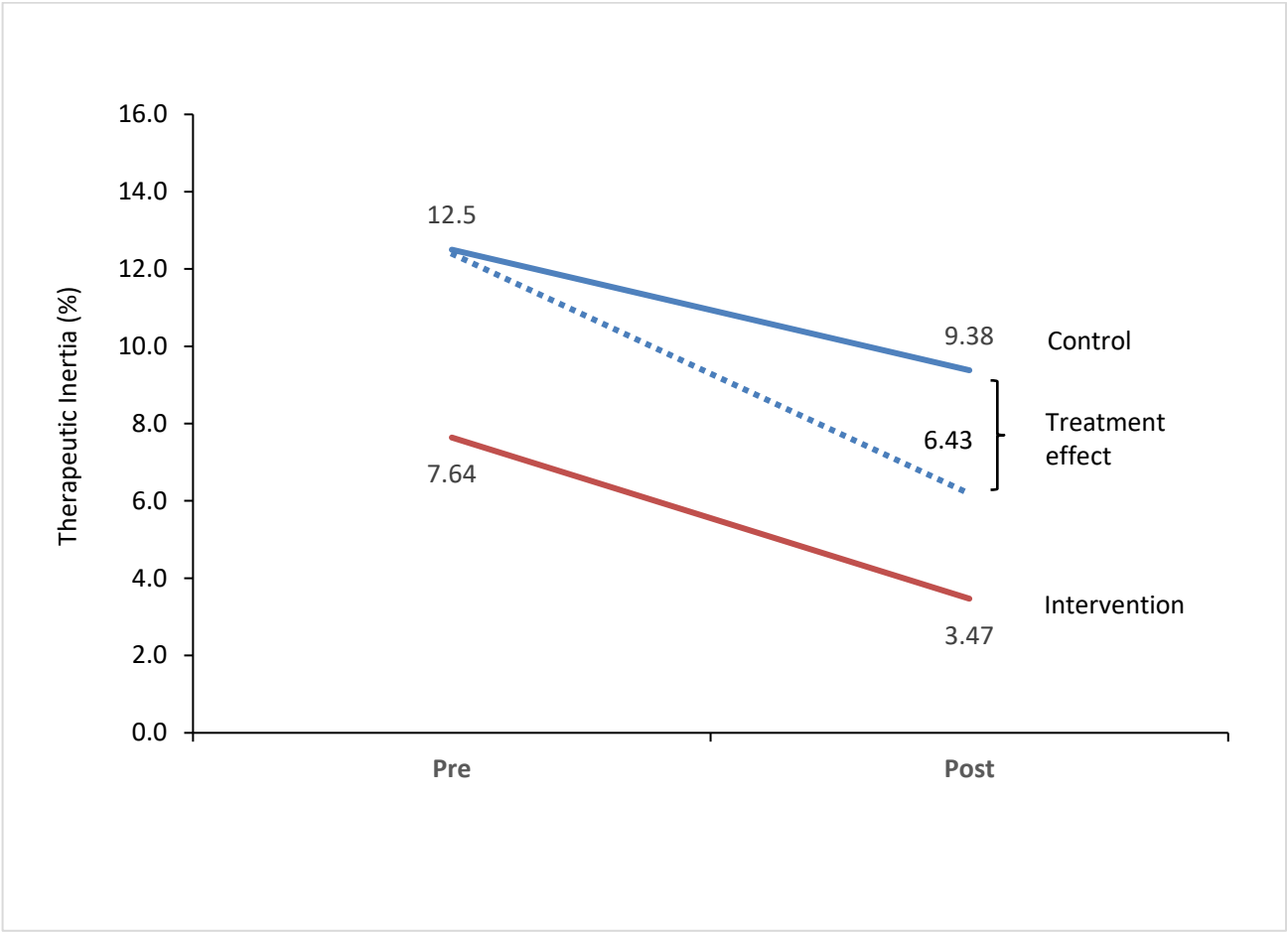


**Figure 5A: Schematic representation of generic mediation analysis**

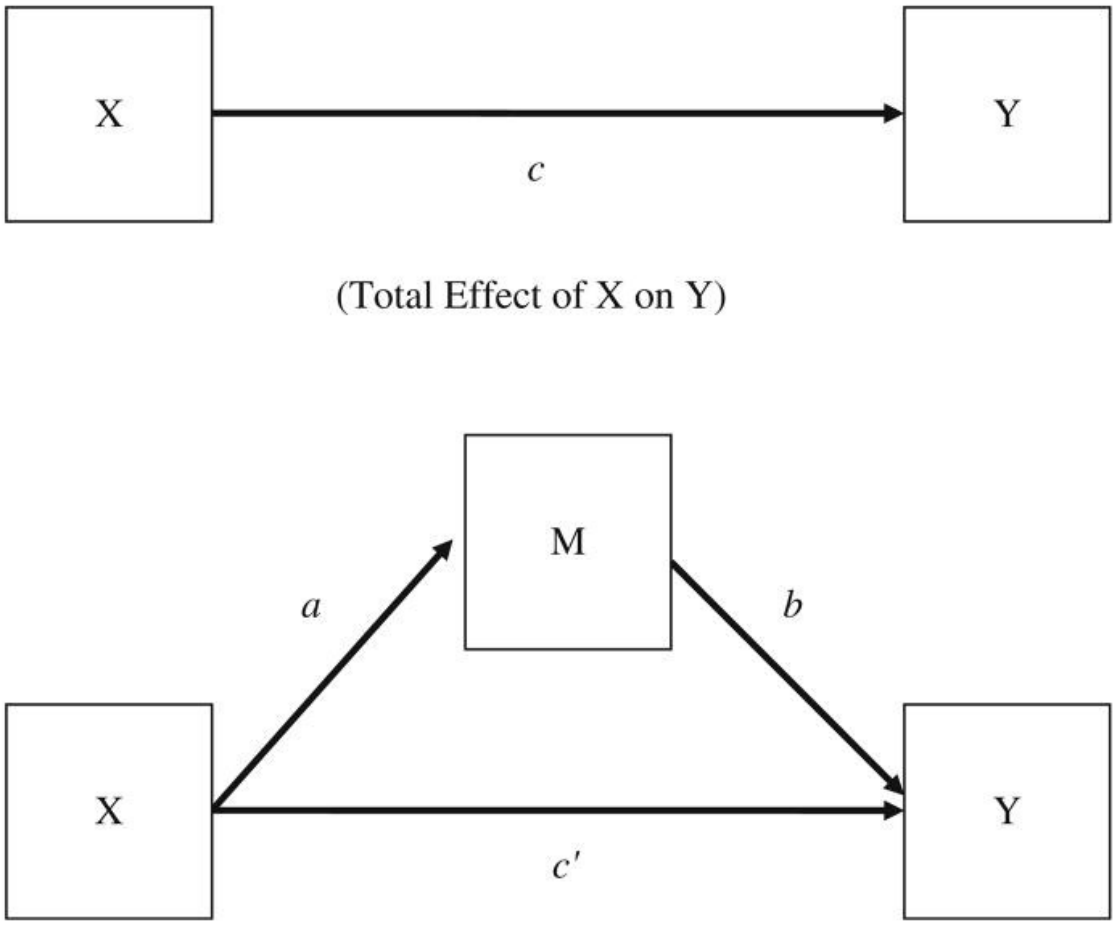


**Figure 5B: Structural equation model used for mediation analysis**

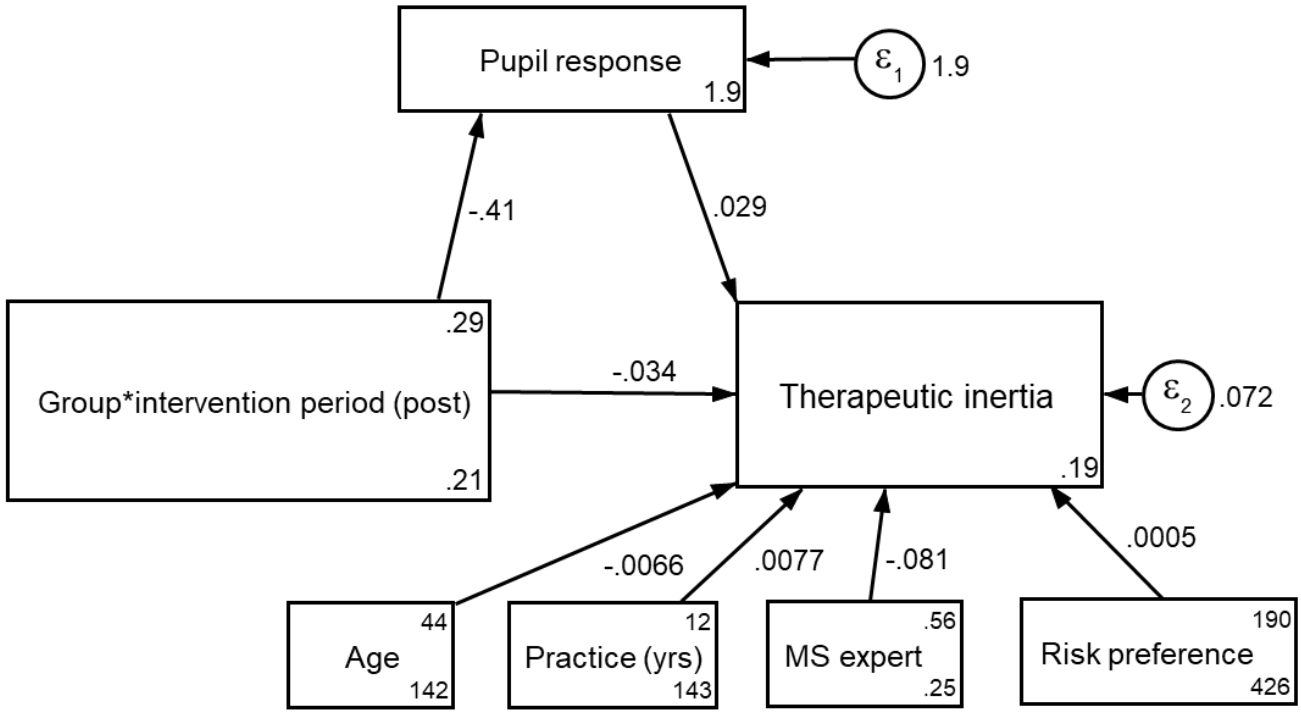
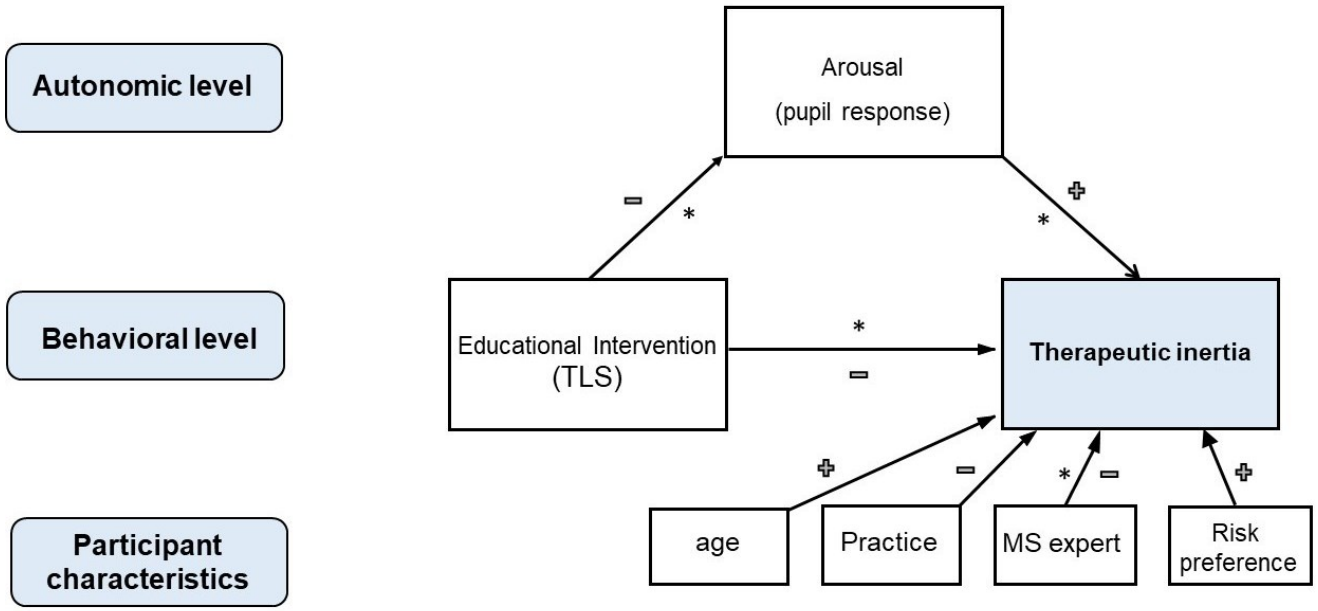


**Figure 5C: Proposed pathways associated with therapeutic inertia**




* p-value for the pathway <0.01

# G. Appendix to Study 7

Multiple Sclerosis and Related Disorders 34 (2019) 17–28

Contents lists available at ScienceDirect

Multiple Sclerosis and Related Disorders

journal homepage: www.elsevier.com/locate/msard

Clinical trial

## Emotional expressions associated with therapeutic inertia in multiple sclerosis care

Gustavo Saposnik[a,b,c,d,*], Jiwon Oh[a,d], Maria A. Terzaghi[a,d], Pawel Kostyrko[a], Fabien Bakdache[e], Alonso Montoya[e], Blessing N.R. Jaja[c], Rosane Nisenbaum[d,f], Christian C. Ruff[b], Philippe N. Tobler[b]

[a] Decision Neuroscience Unit, Division of Neurology, Department of Medicine, St. Michael's Hospital, University of Toronto, 55 Queen St E, Toronto, Ontario M5C 1R6, Canada
[b] Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Switzerland
[c] Decision Neuroscience Unit, St Michael's Hospital, University of Toronto, Canada
[d] Li Ka Shing Knowledge Institute, St. Michael's Hospital, University of Toronto, Canada
[e] Medical Affairs Neuroscience, Hoffmann-La Roche, Mississauga, Canada
[f] Applied Health Research Centre, Centre for Urban Health Solutions, St Michael's Hospital, Toronto, Canada

ARTICLE INFO

*Search terms:*
Multiple sclerosis
Disease-modifying therapy
Therapeutic inertia
Neuroeconomics
Decision making
Risk

ABSTRACT

*Background:* Emotions play a critical role in our daily decisions. However, it remains unclear how and what sort of emotional expressions are associated with therapeutic decisions in multiple sclerosis (MS) care. Our goal was to evaluate the relationship between emotions and affective states (as captured by muscle facial activity and emotional expressions) and TI amongst neurologists caring for MS patients when making therapeutic decisions. *Methods:* 38 neurologists with expertise in MS were invited to participate in a face-to-face study across Canada. Participants answered questions regarding their clinical practice, aversion to ambiguity, and the management of

This paper is published as Saposnik G, Oh J, Terzaghi MA, Kostyrko P, Bakdache F, Montoya A, Jaja BNR, Nisenbaum R, Ruff CC, Tobler PN. Emotional expressions associated with therapeutic inertia in multiple sclerosis care. Mult Scler Relat Disord.,34:17-28, 2019.

Multiple Sclerosis and Related Disorders 34 (2019) 17–28

Contents lists available at ScienceDirect

Multiple Sclerosis and Related Disorders

journal homepage: www.elsevier.com/locate/msard

Clinical trial

# Emotional expressions associated with therapeutic inertia in multiple sclerosis care

Gustavo Saposnik[a,b,c,d,*], Jiwon Oh[a,d], Maria A. Terzaghi[a,d], Pawel Kostyrko[a], Fabien Bakdache[e], Alonso Montoya[e], Blessing N.R. Jaja[c], Rosane Nisenbaum[d,f], Christian C. Ruff[b], Philippe N. Tobler[b]

[a] Decision Neuroscience Unit, Division of Neurology, Department of Medicine, St. Michael's Hospital, University of Toronto, 55 Queen St E, Toronto, Ontario M5C 1R6, Canada
[b] Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Switzerland
[c] Decision Neuroscience Unit, St Michael's Hospital, University of Toronto, Canada
[d] Li Ka Shing Knowledge Institute, St. Michael's Hospital, University of Toronto, Canada
[e] Medical Affairs Neuroscience, Hoffmann-La Roche, Mississauga, Canada
[f] Applied Health Research Centre, Centre for Urban Health Solutions, St Michael's Hospital, Toronto, Canada

ARTICLE INFO

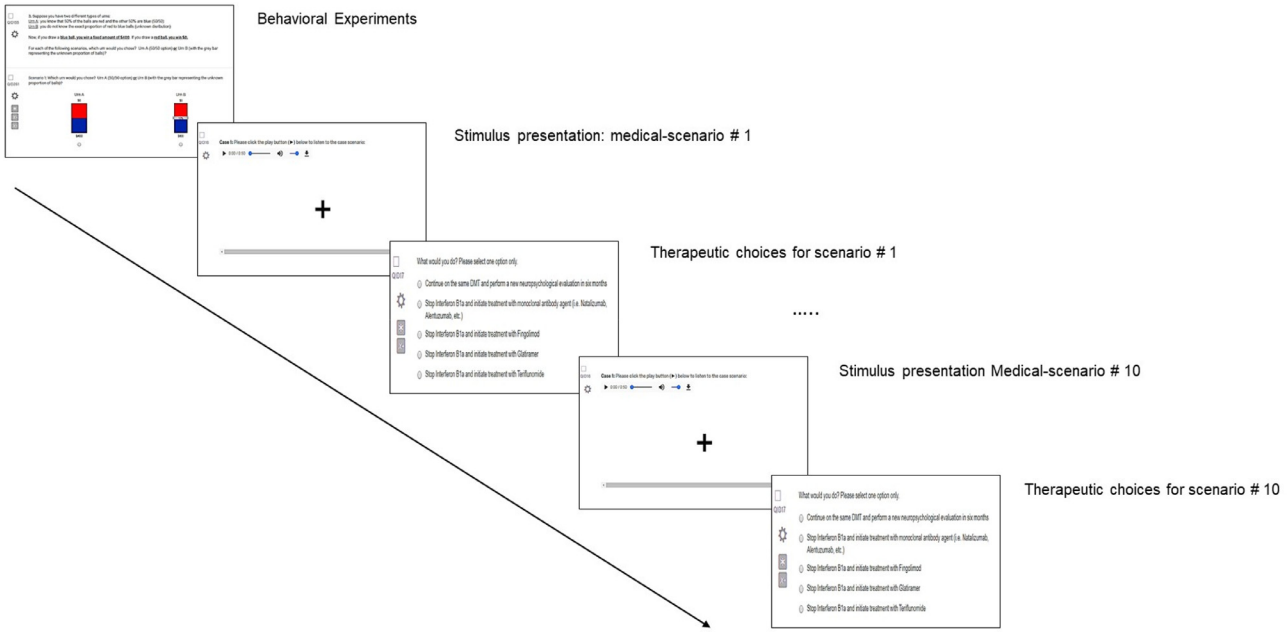
*Search terms:*
Multiple sclerosis
Disease-modifying therapy
Therapeutic inertia
Neuroeconomics
Decision making
Risk

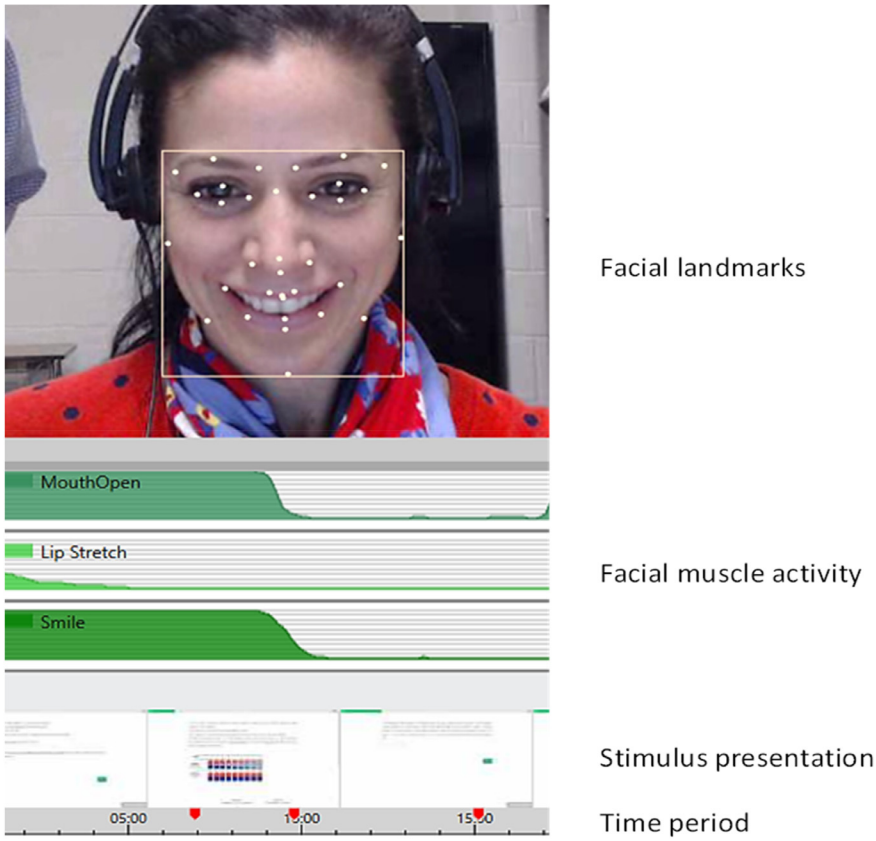
ABSTRACT

*Background:* Emotions play a critical role in our daily decisions. However, it remains unclear how and what sort of emotional expressions are associated with therapeutic decisions in multiple sclerosis (MS) care. Our goal was to evaluate the relationship between emotions and affective states (as captured by muscle facial activity and emotional expressions) and TI amongst neurologists caring for MS patients when making therapeutic decisions.
*Methods:* 38 neurologists with expertise in MS were invited to participate in a face-to-face study across Canada. Participants answered questions regarding their clinical practice, aversion to ambiguity, and the management of 10 simulated case-scenarios. TI was defined as lack of treatment initiation or escalation when there was clear evidence of clinical and radiological disease activity. We recorded facial muscle activations and their associated emotional expressions during the study, while participants made therapeutic choices. We used a validated machine learning algorithm of the AFFDEX software to code for facial muscle activations and a predefined mapping to emotional expressions (disgust, fear, surprise, etc.). Mixed effects models and mediation analyses were used to evaluate the relationship between ambiguity aversion, facial muscle activity/emotional expressions and TI measured as a binary variable and a continuous score.
*Results:* 34 (89.4%) neurologists completed the study. The mean age [standard deviation (SD)] was 44.6 (11.5) years; 38.3% were female and 58.8% self-identified as MS specialists. Overall, 17 (50%) participants showed TI in at least one case-scenario and the mean (SD) TI score was 0.74 (0.90). Nineteen (55.9%) participants had aversion to ambiguity in the financial domain. The multivariate analysis adjusted for age, sex and MS expertise showed that aversion to ambiguity in the financial domain (OR 1.56, 95%CI 1.32–1.86) was associated with TI. Most common muscle activations included mouth open (23.4%), brow furrow (20.9%), brow raise (17.6%), and eye widening (13.1%). Most common emotional expressions included fear (5.1%), disgust (3.2%), sadness (2.9%), and surprise (2.8%). After adjustment for age, sex, and physicians' expertise, the multivariate analysis revealed that brow furrow (OR 1.04; 95%CI 1.003–1.09) and lip suck (OR 1.06; 95%CI 1.01–1.11) were associated with an increase in TI prevalence, whereas upper lip raise (OR 0.30; 95%CI 0.15–0.59), and chin raise (OR 0.90; 95%CI 0.83–0.98) were associated with lower likelihood of TI. Disgust and surprise were associated with a lower TI score (disgust: $p < 0.001$; surprise: $p = 0.008$) and lower prevalence of TI ($OR_{disgust}$: 0.14, 95%CI 0.03–0.65; $OR_{surprise}$: 0.66, 94%CI 0.47–0.92) after adjusting for covariates. The mediation analysis showed that brow furrow was a partial mediator explaining 21.2% (95%CI 14.9%-38.9%) of the association between aversion to ambiguity and TI score, followed by nose wrinkle 12.8% (95%CI 8.9%-23.4%). Similarly, disgust was the single emotional expression (partial mediator) that attenuated (-13.2%, 95%CI -9.2% to -24.3%) the effect of aversion to ambiguity on TI.
*Conclusions:* TI was observed in half of participants in at least one case-scenario. Our data suggest that facial

* Corresponding author at: Division of Neurology, Department of Medicine, St. Michael's Hospital, University of Toronto, 55 Queen St E, Toronto, Ontario M5C 1R6, Canada
*E-mail address:* saposnikg@smh.ca (G. Saposnik).

https://doi.org/10.1016/j.msard.2019.05.029
Received 31 March 2019; Received in revised form 28 May 2019; Accepted 31 May 2019




Facial landmarks

Facial muscle activity

Stimulus presentation

Time period

**Fig. 1.** Study design and facial landmarks. (A) Sequence of study events. After answering demographic and practice-based questions and determining medical and financial ambiguity aversion, participants listened to a case-scenario and then viewed 6 therapeutic choices. This procedure was repeated for each of the 10 case-scenarios ranging from 25 to 50 s. All the stimuli remained on the screen until the participant selected one of the therapeutic choices. Then, the participant was able to see the next screen and play the next case-scenario. The dots between the screen presenting the therapeutic choices #1 and the case-scenario #10 represent the progression through scenarios #2-#9. (B) Facial landmarks. The region of interest in AFFDEX software contains the whole face including eyes, mouth and nose. Each of the 34 facial landmarks are the main unit of study to represent 20 facial expression metrics that are mapped to represent emotional expressions (https://developer.affectiva.com/mapping-expressions-to-emotions/). This figure illustrates the representation of the data at a particular time-point of the study, including: facial metrics, screens of the stimulus presentation, and time landmarks according to the study design and flow. A value of zero indicates no evidence and a value of one the highest evidence that a certain facial metric or emotion is fully expressed. (Levy et al., 2010) .

metrics (e.g. brow furrow, nose wrinkle) and emotional expressions (e.g. disgust) are associated with physicians' choices and partially mediate the effect of aversion to ambiguity on TI.

## 1. Background

The role of emotions in decision-making has been investigated for decades. Recent studies have shown that emotions are the dominant driver of the majority of meaningful goal-directed decisions in life (Ekman, 2007; d'Acremont and Bossaerts, 2012). Different emotions (fear, disgust, stress, surprise, etc.) manifested by facial muscle activation can modulate our perceptions and valuation of individual choices by activating different pathways involving the striatum, orbitofrontal cortex, the medial and dorsolateral prefrontal cortex, the inferior parietal cortex, the amygdala and/or insular cortex (Phelps et al., 2014; Cohen, 2005; Ekman and Friesen, 2003).

Previous studies have shown that decision making was associated with muscle activation (e.g. brow furrow, brow raise, lip pucker, mouth opening) and emotions (e.g. fear, sadness, anger, surprise) in consumers or healthy volunteers (Phelps et al., 2014; Lerner et al., 2003, 2015). For example, fear appears to be associated with pessimistic risk assessments and risk-averse choices, whereas anger can provoke an optimistic estimations of risk and risk-seeking behavior (Lerner et al., 2003, 2015; Kligyte et al., 2013) Similarly, some emotions, such as anger, surprise, and optimism are associated with participants' tolerance to ambiguity and the selection of optimal choices (Chesney and Reiter, 2016; FeldmanHall et al., 2016). However, limited information is available regarding how facial muscle activity (and derived emotional expressions) relate to physicians' therapeutic decisions.

Therapeutic inertia (TI) is a term that was introduced in 2006 to define the absence of treatment initiation or intensification when treatment goals are unmet (O'Connor et al., 2005; Mohan and Phillips, 2011; Okonofua et al., 2006). TI is a common phenomenon affecting 50% to 90% of doctors caring for patients with chronic conditions (e.g. hypertension, diabetes, multiple sclerosis) and leading to poorer clinical outcomes and higher health care costs (O'Connor et al., 2005; Mohan and Phillips, 2011; Okonofua et al., 2006; Burks et al., 2017). Previous studies have identified factors associated with TI (Cooke et al., 2012; Saposnik and Montalban, 2018), and physician factors (e.g. aversion to ambiguity) are considered the main contributors (Saposnik and Montalban, 2018; Lebeau et al., 2014; Saposnik et al., 2017a). To our knowledge, there are no data showing a relationship between facial muscle activation, emotional expressions, and therapeutic decisions under uncertainty (or ambiguity) among practicing physicians.

In this study, we evaluated facial muscle activation (and emotional expression) associated with therapeutic choices, particularly TI. We also sought to evaluate the mediation effect between a physical (e.g. facial muscle activity) or emotional (fear, disgust, surprise) response with a therapeutic decision. Given the known associations between specific facial muscle activation and emotional expression (anger, fear, disgust, surprise, etc.) with an increased attention response that precedes participants' choices (FeldmanHall et al., 2016; Stöckli et al., 2018; McDuff et al., 2017), we hypothesized that facial muscle activity (e.g. upper lip raise) and emotional expression (disgust, surprise) would increase participants' awareness and therefore mediating the relationship between aversion to ambiguity and TI. We assessed emotional expressions amongst physicians who care for people living with multiple sclerosis (MS) as this care model is representative of the paradigm of complex therapeutic decisions (e.g. multiple therapeutic options with a broad therapeutic range- e.g. different safety and efficacy profiles) in the management of a chronic medical condition.

## 2. Methods

### 2.1. Study design

We conducted a cross-sectional study using the online platform Qualtrics. The study included 10 MS case-vignettes to evaluate TI and 2 behavioural experiments to determine subject's attitudes towards ambiguity. Case-scenarios were designed by our research team and MS experts (JO, GS). Overall, 8 cases aimed to assess appropriate escalation of treatment (whereby an absence of treatment change corresponds to TI), while the remaining 2 cases were designed as controls (no indication for treatment escalation as there was no evidence of a clinical relapse and disease activity on brain imaging). After completing demographic information and questions regarding their current clinical practices, participants were exposed to behavioral experiments assessing ambiguity aversion and then responded to case-scenarios (Fig. 1).

Behavioural experiments were designed to assess aversion to ambiguity in the health and financial domains as previously reported by our group (Saposnik et al., 2017a, 2016a; Anderson and Mellor, 2008). Ambiguity aversion is defined as dislike for events with unknown probability over events with known probability (Levy et al., 2010). For example, an ambiguity-averse individual would rather choose a treatment where the probability of benefits or side effects are known (even if these are somewhat unfavourable) over one where these probabilities are unknown. Specifically, in the health domain, participants were asked to choose between Treatment A (50% probability of survival) or "Treatment B" (the probability of survival is unknown). In the financial domain, participants were asked to choose between a visual option with known 50/50 probability of winning 400 or 0 US$ versus an option with unknown probability of the same outcomes. In both domains, we used grey bars to represent five levels (10%, 30%, 50%, 70%, and 90%) of unknown probability. Aversion to ambiguity was indexed in two-ways: binary (preference for the known probability in all 5 levels) and as continuous variable (number of levels that participants selected the known probability over 5). Details of the protocol and case-scenarios were previously published (Saposnik et al., 2017a, 2016a, 2018)

*Participants:* Practicing neurologists actively involved in the care of people living with MS from across Canada were invited to participate in our study by the Canadian Network of MS Clinics and Neuro-sens (Neuro-sens.com). These networks capture most of these neurologists in Canada. Participants were recruited from December 13, 2017 to March 2, 2018. Physicians whose practice focuses primarily on caring for MS patients were classified as 'MS specialists'.

The study was conducted in an ambulatory clinic-type setting to mimic the current clinical environment. Room temperature, light conditions, and participants' sitting positions were standardized. We used a high definition webcam (Logitech Pro 920©) to capture facial movements. All participants had at least 90% muscle detection by the camera during the study period. Facial detection algorithms from AFFDEX (see below) were integrated with the Qualtrics survey platform through iMotions software (iMotions.com). The mean (median) duration of the study was 44.9 (39.9) minutes. Participants were compensated with 400 Canadian dollars. Written informed consent was obtained from all participants. The study was approved by the Research Ethics Board of St. Michael's Hospital, University of Toronto, Canada.

### 2.2. Assessment of emotional expressions

We used AFFDEX, a machine learning algorithm software that detects for emotional expressions based on facial muscle activity

(Stöckli et al., 2018; iMotions 2016). AFFDEX has been validated in more than 6 million facial videos from over 87 countries showing an excellent accuracy (area under the curve greater than 0.9) (https://www.affectiva.com/how/how-it-works/, accessed Feb 28, 2019). This algorithm uses different features to identify 34 facial landmarks (e.g. eye corners, eye centers, nose tip, mouth corner) with a threshold area, discarding background regions (Fig. 2). The region of interest (ROI) contains the whole face including eyes, mouth and nose. AFFDEX applies distinct analytical procedures to identify emotional expressions (https://developer.affectiva.com/mapping-expressions-to-emotions/). During our study, facial detection was recorded to analyze each video frame. Eye blinking and closure were filtered-out. AFFDEX uses frames with a positive detection for the subsequent analysis.

Facial muscle activity is the main unit of study in emotional expressions. Facial movements are detected and mapped on probability values of emotional states (e.g. sadness, joy, disgust, anger, surprise, fear, contempt). The probabilities returned by the AFFDEX module range between zero and one. A value of zero indicates no evidence and

**Figure 2. (A) Prevalence of facial muscle activations**

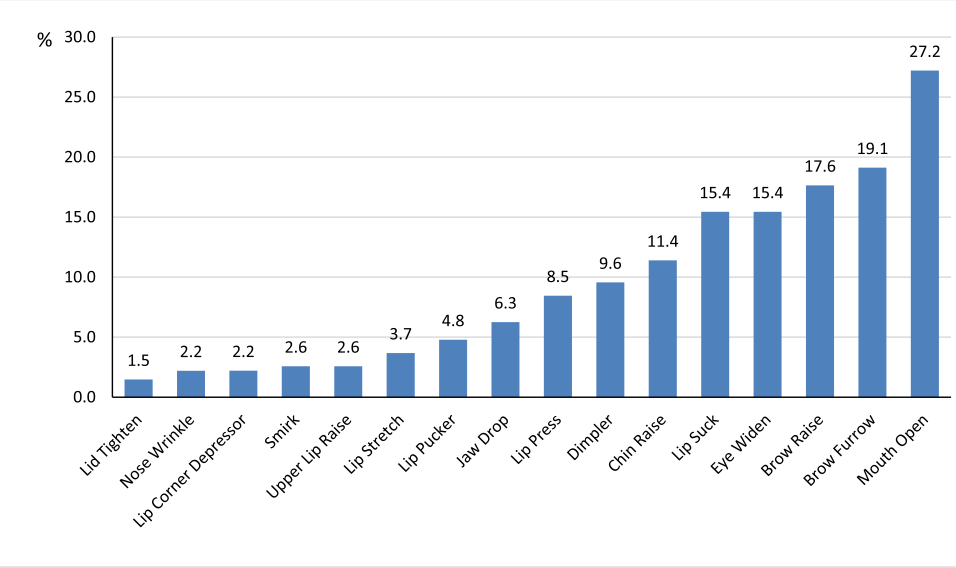



**Figure 2. (B) Prevalence of facial muscle activations by TI status**

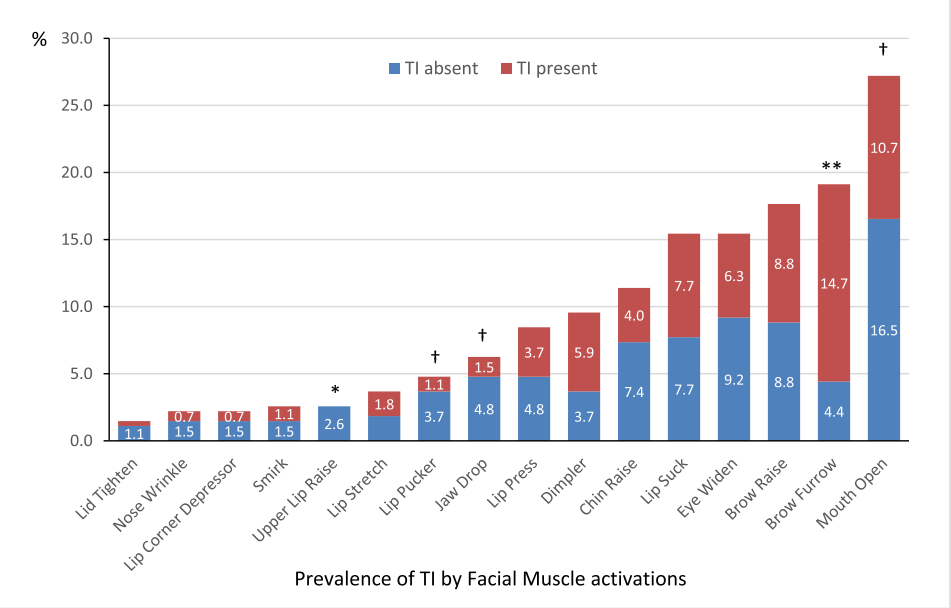
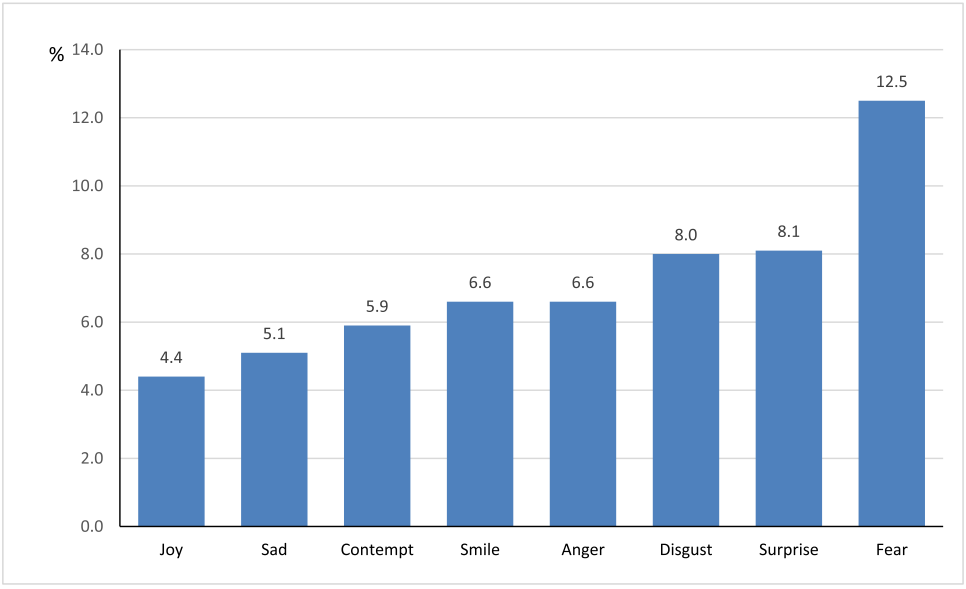



**Fig. 2.** Facial muscle activations overall and in relation to TI status. (A) Overall proportion of facial muscle activations in ascending order. Values at the top of the bars represent the proportion of muscle activation during the study period. (B) Represents the distribution of facial muscle activations as shown in (A) stratified by responses with (red) and without (blue) TI. Values within bars represent the proportion of muscle activations by TI status. For example, brow furrow activation was observed in 19.1% of responses, 14.7% among participants with TI and the remaining 4.4% among participants without TI. The x-axis represents individual facial metric as identified by the AFFDEX software for activations greater than 1%. (C) Overall proportion of emotional expressions in ascending order. Values at the top of the bars represent the proportion of emotional expressions during the study period. (D) Represents the distribution of emotional expressions as shown in (C) stratified by responses with (red) and without (blue) TI. For example, disgust was observed in 8% of responses, 5.8% among participants without TI and the remaining 2.2% among participants with TI. The x-axes represent each individual emotional expression as mapped by the AFFDEX software.

208

* indicates *p*-values < 0.01, ** indicates *p*-values < 0.001, † indicates *p*-values <0.05->0.01 for differences between participants with and without TI.

**(C) Prevalence of emotional expressions overall**




**(D) Prevalence of emotional expressions by TI status**

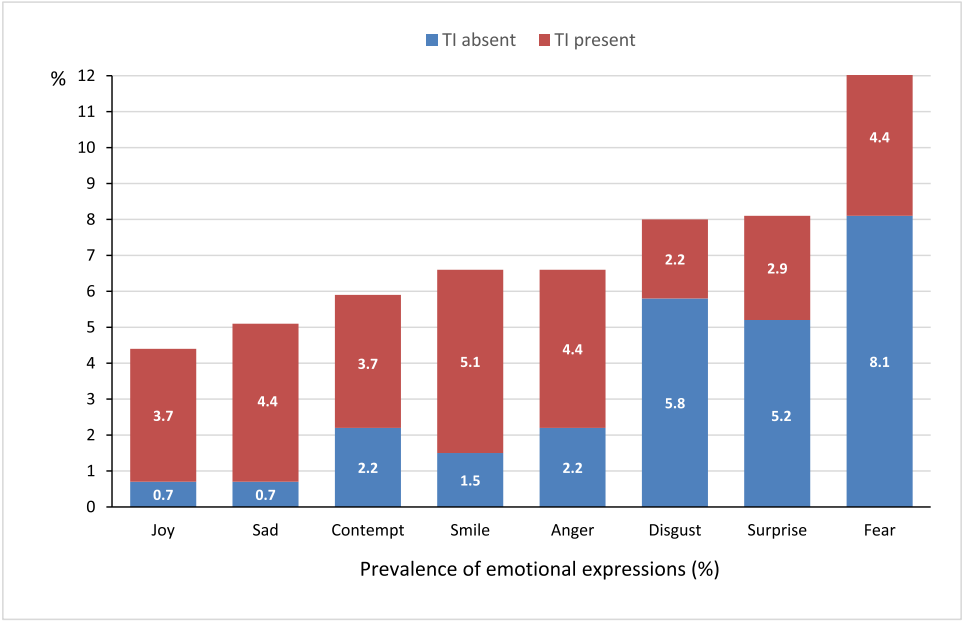



**Fig. 2.** (*continued*)

a value of one the highest evidence that a certain emotion is fully expressed (iMotions 2016). We used raw values of each individual's facial expression to directly compare amongst participants. This approach mitigates potential errors in the algorithms created to represent emotional expressions due to lack of matching with pre-defined facial muscle activity.

We use a proxy measure of participants' arousal by combining the level of attention (a summary measure of the time frame each participant was looking at the screen) and engagement (a weighted sum of facial expressions). We compared facial muscle activity and emotional expressions between participants with and without TI.

*Multiple sclerosis and definitions*: In the context of MS, TI is defined as the lack of treatment initiation or escalation when there is evidence of disease activity, based on clinical evidence and neuroimaging markers (Saposnik et al., 2016a; Freedman et al., 2018; Freedman et al., 2013). A more proactive management strategy, including earlier use of high-efficacy DMTs and close monitoring of the clinical and radiological response to treatment is recommended to slow the progression of physical and cognitive impairments in patients with relapsing-remitting multiple sclerosis (RRMS) (Noyes and Weinstock-Guttman, 2013; Sormani et al., 2013; Duquette et al., 2016). Early treatment escalation has been shown to reduce relapse rates, disability progression, and MRI activity (Prosperini et al., 2012; Harding et al., 2019). For the primary analysis, we used an accepted definition of disease activity that would prompt treatment initiation or escalation (Freedman et al., 2018; Prosperini et al., 2014; Bermel et al., 2013). Disease activity was defined as the presence of a clinical relapse plus the presence of more than four new brain lesions in follow-up magnetic resonance imaging (MRI) scans or at least one gadolinium-enhancing lesion (Prosperini et al., 2014; Bermel et al., 2013).

TI was measured as both a continuous score and as a binary variable. The TI score corresponded to the number of case-scenarios where

treatment initiation or escalation was warranted but not provided (numerator) divided by the total number of case-scenarios where TI could occur (denominator; $n = 8$). TI as a binary variable (presence/absence) was determined as the lack of treatment initiation or escalation given disease activity in at least one case-scenario.

*Outcome measures:* The primary outcomes of the study was the association between facial muscle activity and inferred emotional expression of participants at stimulus presentation (audio introducing MS case-scenarios under uncertainty) when making therapeutic choices and TI.

### 2.3. Statistical analysis

We utilized two analytical approaches: (i) a descriptive assessment of facial muscle activation and emotional expressions, and (ii) a mediation analysis to assess how the association between aversion to ambiguity and TI may be mediated by facial activations and emotional expressions. Mediation analysis is a technique commonly used in social sciences and consumer research to make causal inferences about the influence of specific factors (e.g. demographic variables, participant's characteristics, etc.) on an outcome via a third variable (called 'mediator') (MacKinnon et al., 2007; VanderWeele, 2016). A mediator is a variable that modulates the relationship between that factor with the outcome of interest (MacKinnon et al., 2007; VanderWeele, 2016). In our analysis, the independent variable was aversion to ambiguity and the dependent variable was therapeutic inertia. Facial muscle activity or an emotional expression were individually included as mediators. Further details are illustrated in the appendix (Figs. e1 and e2).

The primary analysis was a descriptive assessment of the presence of facial muscle activation and emotional expression among participants with and without TI (binary) and by the TI score. For each screen face by participants, we calculated the percentage of the frames in which each facial muscle was detected relative to the total number of available frames as part of the AFFDEX software. Then, we identified the time period of the stimulus presentation and the time period of participants' responses when making therapeutic decisions to specifically evaluate the association between facial muscle activation and emotional expressions during these two critical events. Finally, we compared the percentages of facial muscle activation and emotional expressions between participants with and without TI and related them to the TI score. We used a proxy measure of participants' arousal defined as a summary score between attention (range 0–100) and engagement (range 0–100).

Specifically, we used mixed effects logistic and linear models to assess relationships between TI (and TI score) and the percentage of facial muscle movements (and emotional expressions) accounting for clustering (repeated observations on participants). The analysis was adjusted for the following explanatory variables: age, sex, specialist status (MS expert vs. general neurologists). Practice setting (academic vs non-academic), percentage of time devoted to clinical care, and number of MS patients assessed per week had no significant impact on the association between emotional expressions and TI.

We previously found an association between aversion to ambiguity and increased prevalence of TI (Saposnik et al., 2017a). Here, we aimed to replicate this association and evaluated whether this observed association is mediated by facial muscle activation or emotional expression. For the mediation analysis we used the STATA command 'medeff' (see details of the models in the Appendix) (Valeri and Vanderweele, 2013; Fairchild and MacKinnon, 2009). We also use structural equation modeling (SEM) to graphically represent the estimated mediation effects of facial metrics or emotional expressions (see details and interpretation of graphs in the Appendix) (Verkuilen, 2006).

In a sensitivity analysis, we considered the effect of adding participants' number of MS patients seen per week, practice type (academic vs. non-academic), or years of practice instead of participants' expertise in the multivariate models.

Goodness of fit was assessed by the c-statistic for TI (binary outcome) and R-squared for the TI score. All tests were 2-tailed, and $p$-values $<0.05$ were considered significant. We used STATA 13 (College Station, TX: StataCorp LP) to conduct all analyses.

To facilitate the interpretation of findings, we performed the following four analyses:

(1) We evaluated the prevalence of facial muscle activations and emotional expressions; (2) We examined their association with the likelihood of TI and the TI score; (3) We assessed the relationship between facial metrics and emotional expression with ambiguity aversion (main predictor of TI in our previous studies)(18); and (4) we conducted a mediation analysis to determine whether facial muscle activation or emotional expression modulate the relationship between the aversion to ambiguity (independent variable) and TI (outcome).

## 3. Results

### 3.1. Participant characteristics

Of the 38 neurologists who were invited to participate in the emotional recognition study, 34 cooperated (cooperation rate: 89.5%) and 34 (completion rate: 100%) completed the study. The mean age (SD) of study participants was 44.6 ( ± 11.6) years; 13 participants (38.2%) were female. Twenty participants (58.8%) primarily focused their practice on MS care. Participants had on average 12.5 ( ± 12) years of experience and assessed 23.1 ( ± 16) MS patients per week. Table 1 summarizes baseline characteristics of the study population.

TI was present in 50.0% of participants in at least one case-scenario. The mean TI score was 0.74 ( ± 0.90), and the range was 0 to 3.

(1) Prevalence of facial or emotional expressions:

The most commonly observed muscle activations included: mouth open (23.4%), brow furrow (20.9%), brow raise (17.6%), and eye widening (13.1%) (Fig. 2A). Brow furrow was associated with TI ($p < 0.001$). The most commonly decoded emotional expressions included: fear (5.1%), disgust (3.2%), sadness (2.9%), and surprise (2.8%) (Fig. 2C). Differences in facial muscle activation and emotional expressions by TI status are represented in Fig. 2B and D.

Participants with muscle facial activations and emotional expressions had higher arousal scores. For example, arousal scores were significantly higher among participants with disgust (180.7 vs. 133.1; $p = 0.04$), surprise (77.6 vs. 122.4; $p = 0.02$), fear (189.9 vs 131.6; $p = 0.02$). Similar findings were observed for facial muscle activations associated with TI (e.g. brow furrow [$p < 0.001$], brow raise [$p < 0.001$], lip suck [$p < 0.001$], mouth open [$p = 0.02$], nose wrinkle

**Table 1**
Baseline characteristics of participants.

| Characteristics | Total (%) $n = 34$ |
|---|---|
| **Age** (mean ± SD), in years | 44.6 ± 11.6 |
| **Age ≥ 50 years** | 13 (38.2) |
| **Sex** | |
| Female | 13 (38.2) |
| **Specialty** | |
| MS specialists | 20 (58.8) |
| General Neurologists who care for MS patients | 14 (41.2) |
| **Practice setting** | |
| Academic | 28 (82.4) |
| Community | 6 (17.6) |
| **% time in clinical practice** | |
| 50–74% of their time | 16 (47.2) |
| Greater than 75% | 15 (44.1) |
| **Years in practice** (mean ± SD) | 12.5 ± 11.8 |
| **MS patients seen per week** (mean ± SD) | 23.1 ± 15.8 |
| **Author of a peer-reviewed publication in the last 12 months** | 22 (64.7) |

Numbers in brackets indicate percentages.

[$p < 0.01$]).

(1) Basic associations between TI and facial or emotional expressions:

The multivariate mixed effects logistic regression after adjustment for age, sex, and physicians' expertise revealed that brow furrow (OR 1.04; 95%CI 1.003–1.09) and lip suck (OR 1.06; 95%CI 1.01–1.11) were associated with an increase in TI prevalence, whereas upper lip raise (OR 0.30; 95%CI 0.15–0.59), chin raise (OR 0.90; 95%CI 0.83–0.98), and nose wrinkle (OR 0.08; 95%CI 0.007–0.97) were associated with lower likelihood of TI (c-statistic: 0.889). Similar findings were obtained with linear mixed models (brow furrow: $p = 0.05$; lip suck: $p < 0.001$; nose wrinkle: $p = 0.017$, upper lip raise: $p < 0.001$; chin raise: $p < 0.001$; R-squared: 0.373) where the TI score as the outcome of interest.

In the emotional expression analysis, the presence of disgust (characterized by nose-wrinkle and upper lip raise) and surprise (characterized increased brow raise and decrease brow furrow) were associated with lower prevalence of TI scores (disgust: $p < 0.001$; surprise: $p = 0.008$) and TI ($OR_{disgust}$: 0.14, 95%CI 0.03–0.65; $OR_{surprise}$: 0.66, 94%CI 0.47–0.92) after adjusting for age, sex and physicians' expertise. Fear was not associated with either TI ($OR_{fear}$: 0.37, 95%CI 0.03–5.43) or the TI score ($p = 0.14$).

(1) Relation between facial and emotional expressions to ambiguity aversion:

In our previous studies, aversion to ambiguity in the financial domain was the most relevant predictor of TI (Saposnik et al., 2017a). Similarly, in the present study, 19 (55.9%) participants never chose an ambiguous alternative in the financial domain and 11 (32.4%) in the health domain. The multivariate analysis adjusted for age, sex and MS expertise showed that aversion to ambiguity in the financial (OR 1.56, 95%CI 1.32–1.86) and health (OR 1.12, 95%CI 1.02–1.22) domains were independent predictors of TI (Table 2). Similarly, for every 20% increase in the degree of ambiguity, there was a 21.5% increment (95%CI 3.0%−40.0%) in the TI score.

Given the consistent association between aversion to ambiguity and TI in this study and prior studies, we also explored the association between facial muscle activity and emotional expression and ambiguity aversion. The multivariate analysis revealed that mouth opening (OR 2.10, 95%CI 1.35–3.26; $p = 0.001$), brow furrow (OR 2.93, 95%CI 1.84–4.65; $p = < 0.001$), chin raise (OR 3.16, 95%CI 1.51–6.62; $p = 0.002$), and lip suck (OR 0.38, 95%CI 0.21–0.70; $p = 0.002$) were the facial muscle activations associated with higher aversion to ambiguity in the financial domain. Disgust (OR 0.22, 95%CI 0.08–0.65; $p = 0.006$) was the single emotional expression associated with lower aversion to ambiguity.

The mixed linear regression analysis adjusted for age, sex, participants' expertise, emotional expression (disgust) or facial muscle activation (brow furrow, mouth opening, and lip suck) and degree of aversion to ambiguity in the financial domain is presented in Fig. 3. For every 20% increase in the degree of ambiguity (e.g. from 10% to 30%, 30% to 50% and so on), there was a 21.5% increment (95%CI 16.9%−26.0%) in the TI score (R-squared 0.35) (Fig. 3A). Nearly identical results were observed when the linear mixed model included facial muscle activations instead of emotional expression (R-squared 0.38) (Fig. 3B). In contrast, there were no associations between facial muscle activation and emotional expression with aversion to ambiguity in the health domain.

(1) Mediation analysis: facial and emotional expressions modulate the relationship between aversion to ambiguity and TI

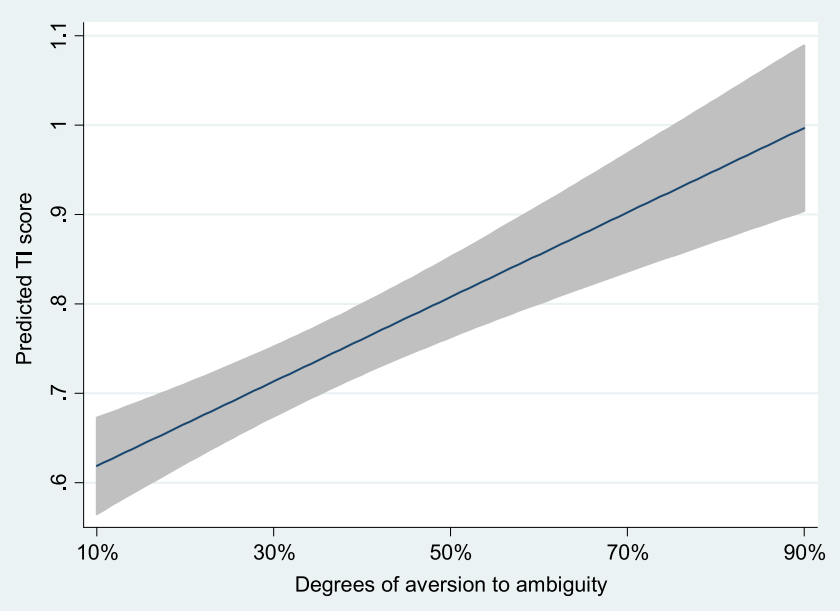
We found that brow furrow followed by nose wrinkle were the strongest mediators, respectively explaining 21.2% (95%CI

**Table 2**
Relationship between aversion to ambiguity and therapeutic inertia.

| | N of individual responses<br>*n* individual responses among participants with aversion to ambiguity /total number of responses (%) | TI score *<br>$\beta$ coefficient (95%CI) | TI †(present/absent)<br>OR (95%CI) | Area under the curve |
|---|---|---|---|---|
| Aversion to Ambiguity (Financial domain) | 380/680 (55.9) | 0.21 (0.16–0.26) | 1.56 (1.32–1.86) | 0.82 |
| Aversion to Ambiguity (Health domain) | 460/680 (66.7) | 0.06 (0.03–0.08) | 1.12 (1.02–1.22) | 0.78 |

* Mixed linear regression models for TI score (dependent variable) adjusted for age, sex, MS expert and aversion to ambiguity (independent variables).
† Mixed logistic regression models for TI (dependent variable) adjusted for age, sex, MS expert and aversion to ambiguity (independent variables).

**Panel A: Accounting for the effect of emotional expression (disgust)**




**Panel B: accounting for the effect of muscle activations**

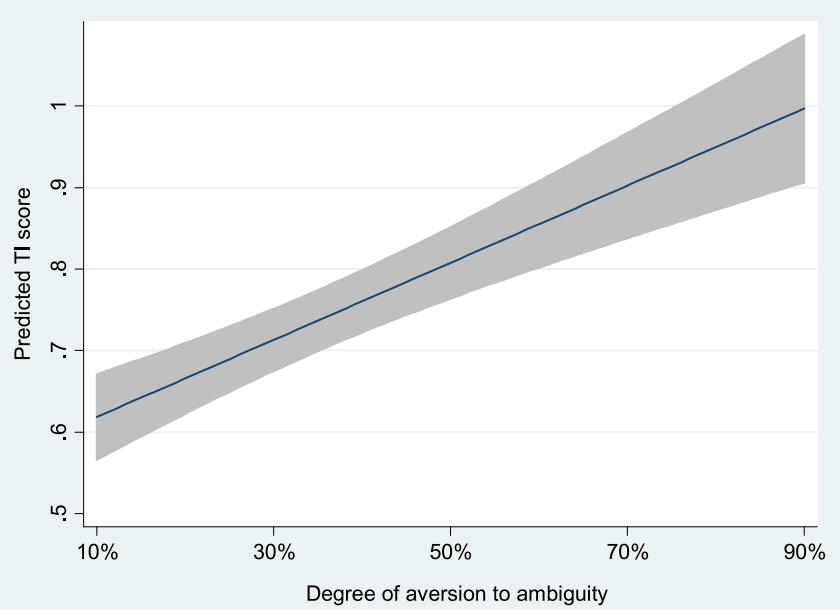



**Fig. 3.** Predicted TI score as a function of degrees of aversion to ambiguity

The mixed linear regression models were adjusted for age, sex, participants' expertise (MS expert vs. general neurologist), disgust (panel A) or facial muscle activation (brow furrow, mouth opening, and lip suck) (Panel B). The gray are represents the 95%CI of the predicted TI score.

For the model accounting for emotional expression (disgust) (Panel A), the R-squared of 0.35 represents the proportion of the variability of the TI score explained by the model. For every 20% increase in the degree of aversion to ambiguity (e.g. from 10% to 30% or from 50% to 90%), there was a 21.5% increment in the TI score.

For the model accounting for facial muscle activations (Panel B), the R-squared of 0.38 represents the proportion of the variability of the TI score explained by the model. For every 20% increase in the degree of aversion to ambiguity (e.g. from 10% to 30% or from 50% to 90%), there was a 19.5% increment in the TI score. The gray are represents the 95%CI of the predicted TI score.

14.9%−38.9%) and 12.8% (95%CI 8.9%−23.4%) of the effect of aversion to ambiguity in the financial domain on TI (Figs. 4A and 4B). Similarly, disgust was the single emotional expression that attenuated (−13.2%, 95%CI −9.2% to −24.3%) the effect of aversion to ambiguity in the financial domain on TI (Fig. 4C).

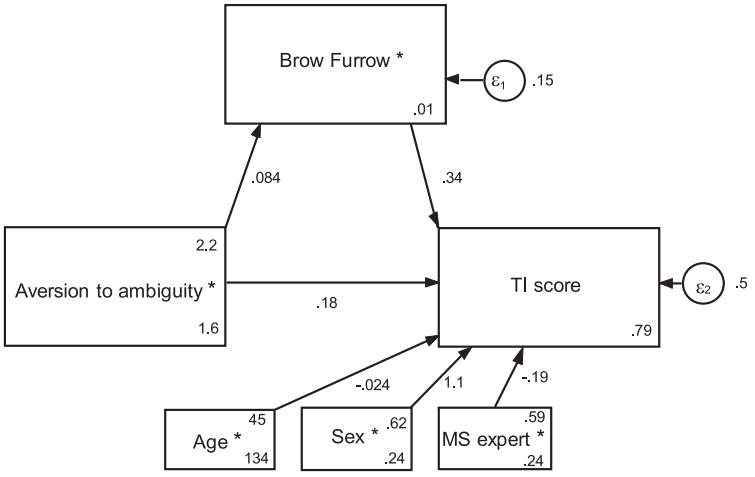
Notably, the direct effect of ambiguity aversion on TI was greater than the indirect effect mediated by brow furrow, nose wrinkle, or disgust (partial mediators). For example, there was a significant, but modest increment (<5%) in the R-square values when adding the facial or emotional variable into the mixed models. This is also reflected in the larger β coefficients for the direct effect between ambiguity aversion and TI compared to the multiplication of β coefficients for the indirect effect (Fig. 4 and figure e2).

Other facial muscle activations (e.g. lip suck: p-value 0.84) and emotional expressions had a non-significant or a negligible effect (e.g. surprise and fear <3%). The sensitivity analysis revealed no changes in the β coefficients for ambiguity aversion when adjusting mixed models for other covariates (see appendix, figures e3 and e4).

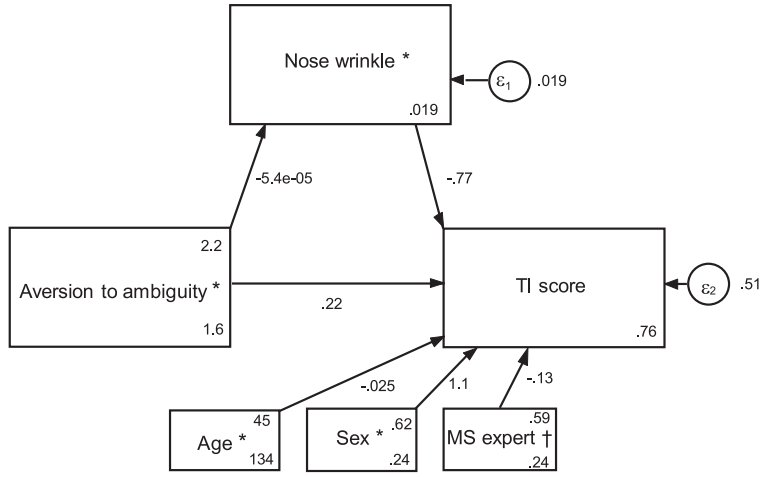
## 4. Discussion

The influence of emotions on the therapeutic decisions of physicians is an important but largely unexplored field. In the present study, we analyzed facial muscle activations and emotional expressions among neurologists while they were making therapeutic decisions. By using the paradigm of complex therapeutic decisions in MS care, we found that emotional expressions (e.g. disgust and surprise) were associated with lower TI. We also observed that facial components of emotional expressions were also associated with TI. Specifically, brow furrow, lip suck and nose wrinkle were associated with an increased prevalence of TI, whereas upper lip raise, and chin raise were associated with a lower likelihood of TI. We also found that aversion to ambiguity increased the

**Panel A. Direct, indirect, and total effect of aversion to ambiguity on TI with brow furrow as mediator**




**Panel B. Direct, indirect, and total effect of aversion to ambiguity on TI with nose wrinkle as mediator**




**Panel C. Direct, indirect, and total effect of aversion to ambiguity on TI with disgust as mediator**

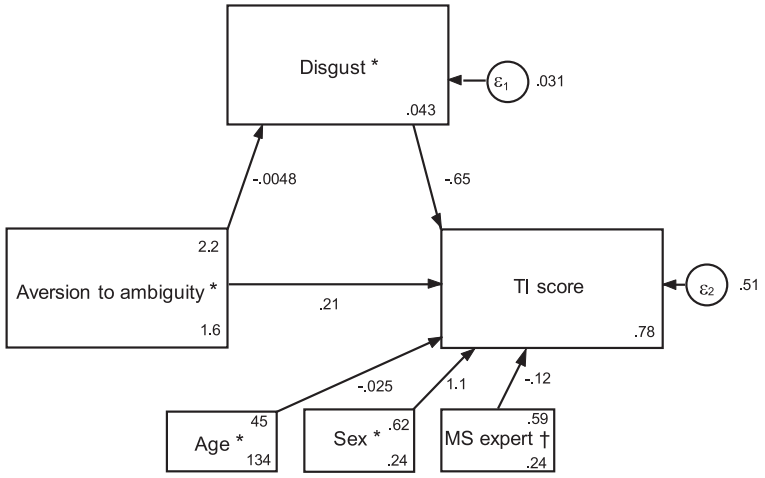



**Fig. 4.** Mediation analysis: graphs derived from structural equation models with a single mediator (see also explanatory figure e1 in the appendix) (A) Structural equation model graph for the modulation of brow furrow (mediator) on the relationship between aversion to ambiguity (independent variable) and therapeutic inertia score (outcome). The R-squared for the model was 0.38. (B) Structural equation model graph for the modulation of nose wrinkle (mediator) on the relationship between aversion to ambiguity (independent variable) and therapeutic inertia score (outcome). The R-squared for the model was 0.34. (C) Structural equation model graph for the modulation of disgust (mediator) on the relationship between aversion to ambiguity (independent variable) and therapeutic inertia score (outcome). The R-squared for the model was 0.35 Age, sex and MS expertise were included as covariates. Values next to the arrows represent β coefficients, ε represent the variance of the mediator and outcome of interest (TI score). Values within each square box represent the mean (upper values) and variance (lower value) of each variable included in the models.

* represents a *p*-value < 0.001, † represents a *p*-value < 0.05 and >0.01 for the total effect models.

213

likelihood of TI. Participants with the aforementioned emotional expressions and muscle facial activations had higher arousal scores compared to those without.

The mediation analysis revealed that disgust was the single emotion that attenuated the effect of aversion to ambiguity in the financial domain on TI. Similarly, the assessment of component processes that mapped to emotional expressions revealed that brow furrow and nose wrinkle were the strongest facial factors explaining 21% and 13% of the influence of aversion to ambiguity on TI.

### 4.1. What is the relevance of our findings for clinical practice?

TI is a common phenomenon affecting 50% to 90% of physicians who manage patients with chronic medical conditions (e.g. hypertension, diabetes, chronic obstructive pulmonary disease, multiple sclerosis, among others) (Cooke et al., 2012; Blasco et al., 2017; Khunti et al., 2017; Ogura and Harada-Shiba, 2016; SisoAlmirall, 2012). TI has been associated with poorer patient outcomes and higher health care costs due to the lack of appropriate treatment escalation (affecting one out of six clinical encounters) leading to higher hospitalizations, greater disability, and lower productivity (Burks et al., 2017; Saposnik et al., 2017a, 2016b; Kobelt et al., 2006). It may occur with insufficient knowledge integration and knowledge-to-action gaps as a result of automatic responses leading to suboptimal therapeutic decisions. Specifically, neurologists caring for MS patients sometimes fail to integrate presented information (e.g. MS severity, relapses within the last three years, imaging findings with the risk of disease progression) with best practice recommendations (Gongora-Ortega et al., 2012; Kennedy et al., 2004; Djulbegovic et al., 2014; Kostopoulou et al., 2017). In the present study, we found an association between facial and emotional expressions with aversion to ambiguity and TI.

Prior studies have demonstrated that interventions increasing physicians' arousal or awareness (e.g. through warning and categorization strategies) were associated with more accurate diagnostic or therapeutic decisions (Mamede et al., 2010, 2017b). A recent randomized clinical trial showed that neurologists who received the traffic light system educational intervention had a 70% reduction in TI (manuscript submitted for publication in Feb 2019). Furthermore, several studies demonstrated a link between specific emotions (e.g. disgust) increasing attention at early stages of visual processing (van Hooff et al., 2014). These prior findings, together with observations from the current study suggest that emotional expressions and strategies that enhance participants awareness (via increasing attention or arousal) may reduce TI.

Furthermore, our findings of (i) brow furrow being associated with both increased TI and ambiguity aversion and (ii) disgust being associated with both reduced TI and lower ambiguity aversion indicates that common emotional factors may contribute to both behaviors. In our previous studies, aversion to ambiguity was the most significant physician-level factor associated with TI (Saposnik and Montalban, 2018; Saposnik et al., 2017a; Reach, 2014) Given the limited training in risk management and formal learning in medical decisions, physicians are clearly vulnerable when handling decisions under uncertainty, especially when having aversion to ambiguity (Dijkstra et al., 2015; Monrouxe et al., 2017; Kostopoulou et al., 2012). Taken together, these findings suggest that interventions reducing TI may partly rely on emotional factors (Saposnik et al., 2017b) and that emotional factors may play a more important role for medical decision making than hitherto assumed.

What brain pathways that may underpin the link between emotions and TI?

Previous studies suggest that the neural mechanisms mediating the relation between affect and decisions depend on a participant's emotional arousal and engagement with the specific choice to be made (Phelps et al., 2014). For example, disgust has been associated with the activation of the insular cortex which may lead to increased arousal modulating the neural responses to aversion to ambiguity, which results in influencing subsequent decision-making (Klucken et al., 2012; Mataix-Cols et al., 2008). Disgust was also shown to increase arousal by modulating emotion-specific attention (van Hooff et al., 2014). This finding is also consistent with an increased arousal score associated with disgust (and its muscle components) in our study.

Traditionally, the striatum, the amygdala, the medial prefrontal, orbitofrontal and insular cortices are thought to process emotional aspects of the decision-making process (Phelps et al., 2014; Lerner et al., 2015). Moreover, the dorsolateral and anterior prefrontal cortices and the posterior parietal cortex may modulate cognitive aspects of decisions (Cohen, 2005). Previous studies showed that stress reduces activity in dorsolateral and orbital parts of prefrontal cortex while it enhances amygdala activity, leading to decreased goal-directed behavior and increased emotional responses (e.g. fear, disgust, contempt) (Otto et al., 2014; Sokol-Hessner et al., 2013). Findings from our study are in

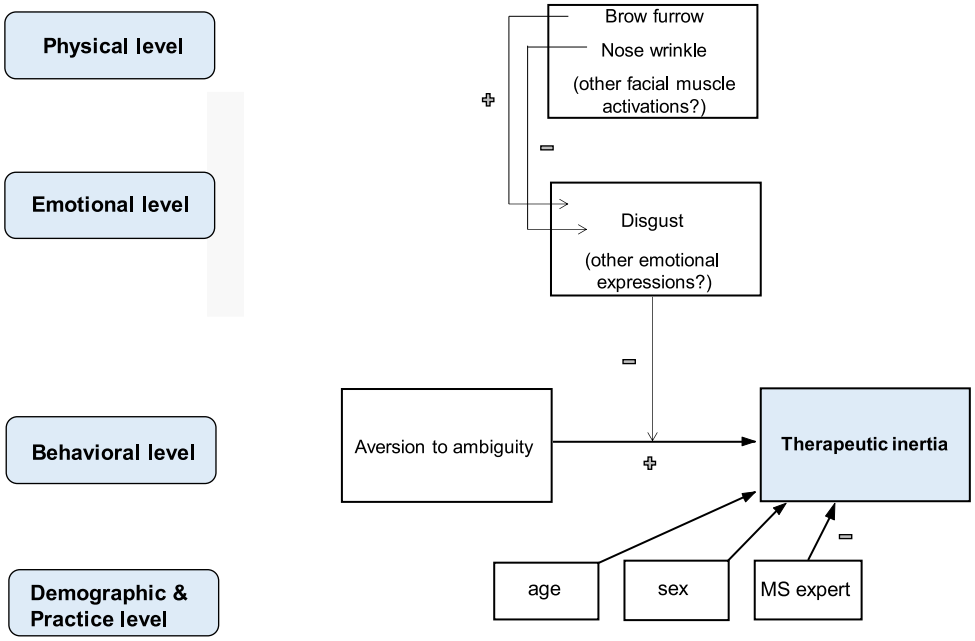



**Fig. 5.** Schematic representation of factors associated with therapeutic inertia (TI)
There is a direct effect of ambiguity aversion and TI and an indirect effect of facial muscle activations (e.g. brow furrow, lip suck) and emotional expressions (e.g. disgust) modulating the relationship between ambiguity aversion and therapeutic inertia. Demographic and practice factors (e.g. number of MS patients seen per week) may also contribute to TI.

214

keeping with this proposed framework as they support an association between facial metrics and emotional expressions (disgust) which may increase participant's awareness/arousal (reflected by increased arousal scores) regarding a compelling decision, thereby showing a reduction in ambiguity aversion and lowering the likelihood of TI (Fig. 5). We could speculate that muscle activations associated with disgust and surprise may reflect increased activity of the insular cortex and amygdala leading to greater arousal and lower TI. Further studies using time-resolved neural methods are needed to test this hypothesis.

## 5. Limitations

Our study has a number of significant limitations. First, although we used a validated software to detect facial muscle activation, the association with emotional expressions may require further assessment. Second, there is a high variability of emotional expressions when participants are exposed to a specific stimulus. As such, our results should be interpreted with caution considering there are no other similar studies available for comparison. Third, the sample size is small affecting the precision of our results (e.g. wider confidence intervals). Fourth, the prevalence of emotional expressions was relatively low likely due to: (i) the strict pre-specified correlation mapping used by AFFDEX software (that combines the concomitant activation of several facial muscles to code for a single emotion) (Stöckli et al., 2018), and (ii) emotions are inherently social, and therefore more neutral expressions are commonly observed when participants are exposed to computer-based simulated scenarios. (de Melo et al., 2014; van 't Wout et al., 2006) Finally, our stimulus was based on case-scenarios that may not truly reflect the therapeutic decisions in clinical practice.

Despite these limitations, our study is the first to show that facial muscle activations and emotional expressions are associated with therapeutic decisions made by physicians who care for MS patients.

## 6. Conclusions

This information helps improve our understanding of the influence of emotional expressions on physicians' therapeutic decisions. These findings, in conjunction with results from a prior study that demonstrated the benefits of an educational intervention on reducing TI have practical clinical implications. With further studies, it may be possible to identify physicians at high risk of having TI by evaluating physical, emotional, and behavioral responses (aversion to ambiguity) and tailor educational interventions to these individuals. Identifying and administering appropriate educational interventions in such situations may facilitate optimal therapeutic decisions in chronic diseases, resulting in better patient outcomes and lower health care costs.

## Authors disclosures

We wish to draw the attention of the Editor to the following facts which may be considered as potential conflicts of interest and to significant financial contributions to this work.

Dr. Gustavo Saposnik is supported by the Heart and Stroke Foundation Career Scientist Awards following an open peer-reviewed competition.

Dr. Alonso Montoya, Fabien Bakdache are employees at the Department of Medical Affairs, Neuroscience, Hoffmann-La Roche Limited, Mississauga, ON, Canada

Maria Terzaghi, Blessing Jaja, Rosane Nisenbaum – no disclosures

Dr. Oh has received research funding from the MS Society of Canada, the National MS Society, Brain Canada, Sanofi-Genzyme, Roche, and Biogen. Dr. Oh has received personal compensation for consulting or speaking from EMD-Serono, Genzyme, Biogen, Novartis, Celgene, and Roche.

Prof. Ruff received research funding from the Swiss National Science Foundation (100019L_173248)

Prof. Philippe Tobler received research funding from the Swiss National Science Foundation (PP00P1_150739 and 100014_165884) and Pfizer.

*Study funding:* The study was sponsored by Division of Neurology at the University of Toronto with the collaboration of Neuro-sens and funded through a research collaboration agreement with Roche Canada. The sponsoring entities were not involved in the design, execution, analysis, and interpretation or reporting of the results. Ocrelizumab (manufactured by Roche) was not included as a therapeutic option in any of the case-scenarios.

## Acknowledgements

The authors are most grateful to all physicians participating in the study.

We appreciate the support from the Department of Economics at the University of Zurich by facilitating access to the Qualtrics platform.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.msard.2019.05.029.

## References

Anderson, L.R., Mellor, J.M., 2008. Predicting health behaviors with an experimental measure of risk preference. J. Health Econ. 27 (5), 1260–1274.
Bermel, R.A., You, X., Foulds, P., Hyde, R., Simon, J.H., Fisher, E., et al., 2013. Predictors of long-term outcome in multiple sclerosis patients treated with interferon beta. Ann. Neurol. 73 (1), 95–103.
Blasco, M., Perez-Martinez, P., Lahoz, C., 2017. Decalogue of the Spanish society of arteriosclerosis to reduce therapeutic inertia. Clin. Investig. Arterioscler. 29 (5), 218–223.
Burks, J., Marshall, T.S., Ye, X., 2017. Adherence to disease-modifying therapies and its impact on relapse, health resource utilization, and costs among patients with multiple sclerosis. ClinicoEcon. Outcomes Res. CEOR 9, 251–260.
Cohen, J.D., 2005. The vulcanization of the human brain: a neural perspective on interactions between cognition and emotion. J. Econ. Perspect. 19 (4), 3–24.
Cooke, C.E., Sidel, M., Belletti, D.A., Fuhlbrigge, A.L., 2012. Review: clinical inertia in the management of chronic obstructive pulmonary disease. COPD 9 (1), 73–80.
d'Acremont, M., Bossaerts, P., 2012. Decision making: how the brain weighs the evidence. Curr. Biol. 22 (18), R808–R810.
Dijkstra, I.S., Pols, J., Remmelts, P., Brand, P.L., 2015. Preparedness for practice: a systematic cross-specialty evaluation of the alignment between postgraduate medical education and independent practice. Med. Teach. 37 (2), 153–161.
Djulbegovic, B., Beckstead, J.W., Elqayam, S., Reljic, T., Hozo, I., Kumar, A., et al., 2014. Evaluation of physicians' cognitive styles. Med. Decis. Mak. 34 (5), 627–637.
Duquette, P., Giacomini, P.S., Bhan, V., Hohol, M., Schecter, R., 2016. Balancing early aggression against risk of progression in multiple sclerosis. Can. J. Neurol. Sci. 43 (1), 33–43.
Ekman, P., Friesen, W.V., 2003. Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues. Ishk.
Ekman, P., 2007. Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life. Macmillan.
Fairchild, A.J., MacKinnon, D.P., 2009. A general model for testing mediation and moderation effects. Prev. Sci. 10 (2), 87–99.
FeldmanHall, O., Glimcher, P., Baker, A.L., Phelps, E.A., 2016. Emotion and decision-making under uncertainty: physiological arousal predicts increased gambling during ambiguity but not risk. J. Exp. Psychol. Gen. 145 (10), 1255–1262.
Freedman, M.S., Selchen, D., Arnold, D.L., Prat, A., Banwell, B., Yeung, M., et al., 2013. Treatment optimization in MS: Canadian MS working group updated recommendations. Can. J. Neurol. Sci. 40 (3), 307–323.
Freedman, M.S., Selchen, D., Prat, A., Giacomini, P.S., 2018. Managing multiple sclerosis: treatment initiation, modification, and sequencing. Can. J. Neurol. Sci. 45 (5), 489–503.
Gongora-Ortega, J., Segovia-Bernal, Y., Valdivia-Martinez Jde, J., Galaviz-DeAnda, J.M., Prado-Aguilar, C.A., 2012. Educational interventions to improve the effectiveness in clinical competence of general practitioners: problem-based versus critical reading-based learning. BMC Med. Educ. 12, 53.
Gordon, N.S., Chesney, S.A., Reiter, K., 2016. Thinking positively: optimism and emotion regulation predict interpretation of ambiguous information. Cogent Psychol. 3 (1), 1195068.
Harding, K., Williams, O., Willis, M., Hrastelj, J., Rimmer, A., Joseph, F., et al., 2019. Clinical outcomes of escalation vs early intensive disease-modifying therapy in patients with multiple sclerosis. JAMA Neurol. 76 (5), 536–541.
iMotions. Facial Expression Analysis. Boston, MA 02111; 2016.
Kennedy, T., Regehr, G., Rosenfield, J., Roberts, S.W., Lingard, L., 2004. Exploring the gap between knowledge and behavior: a qualitative study of clinician action

following an educational intervention. Acad. Med. 79 (5), 386–393.
Khunti, K., Gomes, M.B., Pocock, S., Shestakova, M.V., Pintat, S., Fenici, P., et al., 2017. Therapeutic inertia in the treatment of hyperglycaemia in patients with type 2 diabetes: a systematic review. Diabetes Obes. Metab. 20 (2), 427–437.
Kligyte, V., Connelly, S., Thiel, C., Devenport, L., 2013. The influence of anger, fear, and emotion regulation on ethical decision making. Hum. Perform. 26 (4), 297–326.
Klucken, T., Schweckendiek, J., Koppe, G., Merz, C.J., Kagerer, S., Walter, B., et al., 2012. Neural correlates of disgust- and fear-conditioned responses. Neuroscience 201, 209–218.
Kobelt, G., Berg, J., Lindgren, P., Fredrikson, S., Jönsson, B., 2006. Costs and quality of life of patients with multiple sclerosis in Europe. J. Neurol. Neurosurg. Psychiatry 77 (8), 918–926.
Kostopoulou, O., Russo, J.E., Keenan, G., Delaney, B.C., Douiri, A., 2012. Information distortion in physicians' diagnostic judgments. Med. Decis. Mak. 32 (6), 831–839.
Kostopoulou, O., Sirota, M., Round, T., Samaranayaka, S., Delaney, B.C., 2017. The role of physicians' first impressions in the diagnosis of possible cancers without alarm symptoms. Med. Decis. Mak. 37 (1), 9–16.
Lebeau, J.P., Cadwallader, J.S., Aubin-Auger, I., Mercier, A., Pasquet, T., Rusch, E., et al., 2014. The concept and definition of therapeutic inertia in hypertension in primary care: a qualitative systematic review. BMC Fam. Pract. 15, 130.
Lerner, J.S., Gonzalez, R.M., Small, D.A., Fischhoff, B., 2003. Effects of fear and anger on perceived risks of terrorism: a national field experiment. Psychol. Sci. 14 (2), 144–150.
Lerner, J.S., Li, Y., Valdesolo, P., Kassam, K.S., 2015. Emotion and decision making. Annu. Rev. Psychol. 66 (1), 799–823.
Levy, I., Snell, J., Nelson, A.J., Rustichini, A., Glimcher, P.W., 2010. Neural representation of subjective value under risk and ambiguity. J. Neurophysiol. 103 (2), 1036–1047.
MacKinnon, D.P., Fairchild, A.J., Fritz, M.S., 2007. Mediation Analysis. Annu. Rev. Psychol. 58 (1), 593–614.
Mamede, S., van Gog, T., van den Berge, K., Rikers, R.M., van Saase, J.L., van Guldener, C., et al., 2010. Effect of availability bias and reflective reasoning on diagnostic accuracy among internal medicine residents. JAMA 304 (11), 1198–1203.
Mataix-Cols, D., An, S.K., Lawrence, N.S., Caseras, X., Speckens, A., Giampietro, V., et al., 2008. Individual differences in disgust sensitivity modulate neural responses to aversive/disgusting stimuli. Eur. J. Neurosci. 27 (11), 3050–3058.
McDuff, D., Kodra, E., Kaliouby, R.E., LaFrance, M., 2017. A large-scale analysis of sex differences in facial expressions. PLoS One 12 (4), e0173942.
Melo, C.D., Gratch, J., Carnevale, P.J., 2015. Humans versus computers: Impact of emotion expressions on people's decision making. IEEE Trans. Affect. Comput. 6, 127–136.
Mohan, A.V., Phillips, L.S., 2011. Clinical inertia and uncertainty in medicine. JAMA 306 (4), 383 author reply -4.
Monrouxe, L.V., Grundy, L., Mann, M., John, Z., Panagoulas, E., Bullock, A., et al., 2017. How prepared are UK medical graduates for practice? A rapid review of the literature 2009-2014. BMJ Open 7 (1), e013656.
Noyes, K., Weinstock-Guttman, B., 2013. Impact of diagnosis and early treatment on the course of multiple sclerosis. Am. J. Manag. Care 19 (17 Suppl), s321–s331.
O'Connor, P.J., Sperl-Hillen, J.A.M., Johnson, P.E., Rush, W.A., Biltz, G., 2005. Clinical inertia and outpatient medical errors. In: Henriksen, K, Battles, JB, Marks, ES, Lewin, DI (Eds.), Advances in Patient Safety: From Research to Implementation 2 Advances in Patient Safety, RockvilleMD Concepts and Methodology.
Ogura, M., Harada-Shiba, M., 2016. Clinical inertia in the management of hypercholesterolemia: what clinicians need to do. J. Atheroscler. Thromb. 23 (5), 552–553.
Okonofua, E.C., Simpson, K.N., Jesri, A., Rehman, S.U., Durkalski, V.L., Egan, B.M., 2006. Therapeutic inertia is an impediment to achieving the Healthy People 2010 blood pressure control goals. Hypertension 47 (3), 345–351.
Otto, B., Misra, S., Prasad, A., McRae, K., 2014. Functional overlap of top-down emotion regulation and generation: an fMRI study identifying common neural substrates between cognitive reappraisal and cognitively generated emotions. Cognit. Affect. Behav. Neurosci. 14 (3), 923–938.
Phelps, E.A., Lempert, K.M., Sokol-Hessner, P., 2014. Emotion and decision making: multiple modulatory neural circuits. Annu. Rev. Neurosci. 37 (1), 263–287.
Prosperini, L., Gianni, C., Leonardi, L., De Giglio, L., Borriello, G., Galgani, S., et al., 2012. Escalation to natalizumab or switching among immunomodulators in relapsing multiple sclerosis. Mult. Scler. 18 (1), 64–71.
Prosperini, L., Mancinelli, C.R., De Giglio, L., De Angelis, F., Barletta, V., Pozzilli, C., 2014. Interferon beta failure predicted by EMA criteria or isolated MRI activity in multiple sclerosis. Mult. Scler. 20 (5), 566–576.
Reach, G., 2014. Clinical inertia, uncertainty and individualized guidelines. Diabetes Metab. 40 (4), 241–245.
Saposnik, G., Montalban, X., 2018. Therapeutic inertia in the new landscape of multiple sclerosis care. Front. Neurol. 9, 174.
Saposnik, G., Sempere, A.P., Raptis, R., Prefasi, D., Selchen, D., Maurino, J., 2016a. Decision making under uncertainty, therapeutic inertia, and physicians' risk preferences in the management of multiple sclerosis (DIScUTIR MS). BMC Neurol. 16 (1), 58.
Saposnik, G., Sempere, A.P., Raptis, R., Prefasi, D., Selchen, D., Maurino, J., 2016b. Decision making under uncertainty, therapeutic inertia, and physicians' risk preferences in the management of multiple sclerosis (DIScUTIR MS). BMC Neurol. 16, 58.
Saposnik, G., Sempere, A.P., Prefasi, D., Selchen, D., Ruff, C.C., Maurino, J., et al., 2017a. Decision-making in multiple sclerosis: the role of aversion to ambiguity for therapeutic inertia among neurologists (DIScUTIR MS). Front. Neurol. 8, 65.
Saposnik, G., Maurino, J., Sempere, A.P., Terzaghi, M.A., Ruff, C.C., Mamdani, M., et al., 2017b. Overcoming therapeutic inertia in multiple sclerosis care: a pilot randomized trial applying the traffic light system in medical education. Front. Neurol. 8, 430.
Saposnik, G., Montalban, X., Selchen, D., Terzaghi, M.A., Bakdache, F., Montoya, A., et al., 2018. Therapeutic inertia in multiple sclerosis care: a study of Canadian neurologists. Front. Neurol. 9, 781.
Siso Almirall, A., 2012. Clinical inertia in osteoarthritis. Aten. Prim. 44 (2), 72–73.
Sokol-Hessner, P., Camerer, C.F., Phelps, E.A., 2013. Emotion regulation reduces loss aversion and decreases amygdala responses to losses. Soc. Cognit. Affect. Neurosci. 8 (3), 341–350.
Sormani, M.P., Rio, J., Tintore, M., Signori, A., Li, D., Cornelisse, P., et al., 2013. Scoring treatment response in patients with relapsing multiple sclerosis. Mult. Scler. 19 (5), 605–612.
Stöckli, S., Schulte-Mecklenbeck, M., Borer, S., Samson, A.C., 2018. Facial expression analysis with AFFDEX and FACET: a validation study. Behav. Res. Methods 50 (4), 1446–1460.
Valeri, L., Vanderweele, T.J., 2013. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. Psychol. Methods 18 (2), 137–150.
van 't Wout, M., Kahn, R.S., Sanfey, A.G., Aleman, A., 2006. Affective state and decision-making in the Ultimatum Game. Exp. Brain Res. 169 (4), 564–568.
van Hooff, J.C., van Buuringen, M., El M'rabet, I., de Gier, M., van Zalingen, L., 2014. Disgust-specific modulation of early attention processes. Acta Psychol. 152, 149–157.
VanderWeele, T.J., 2016. Mediation analysis: a practitioner's guide. Annu. Rev. Public Health 37 (1), 17–32.
Verkuilen, J., 2006. Explanatory item response models: a generalized linear and nonlinear approach by P. de Boeck and M. Wilson and generalized latent variable modeling: multilevel, longitudinal and structural equation models by A. Skrondal and S. Rabe-Hesketh. Psychometrika. 71 (2), 415–418.

# Curriculum Vitae

## Personal Details:

First, Last name: Gustavo Saposnik

Date of birth: 14 March 1967

## Education

| | |
|---|---|
| 2015 - 2020 | PhD candidate, Department of Economy, University of Zurich, Zürich, Switzerland, Supervisor(s): Prof. Dr. Christian Ruff and Prof. Dr. Philippe Tobler |
| 2003 - 2006 | Postdoctoral Clinical Stroke Fellow, Stroke Program, Department of Neurology, London Health Sciences Centre, University of Western Ontario, London, Ontario, Canada, Supervisor: Dr. Vladimir Hachinski |
| 2000 - 2003 | Master's in Public Health (MPH), Clinical Effectiveness Program, Joint Program between the University of Buenos Aires & Harvard School of Public Health, Buenos Aires, Capital Federal, Argentina, Supervisor(s): Dr. R. Rubinstein and Dr. Ezequiel Garcia Elorrio |
| 1997 - 1999 | Postdoctoral Clinical Stroke Fellow, Stroke Program, Department of Neurology, Ramos Mejia Hospital, University of Buenos Aires, Buenos Aires, Argentina, Supervisor: Dr. Raul Rey |
| 1993 - 1996 | Neurology Residency, Neurology, Department of Neurological Sciences, University of Buenos Aires, Buenos Aires, Argentina |
| 1991 - 1993 | Internship, Internal Medicine, Dept of Medicine, University of Buenos Aires, Buenos Aires, Buenos Aires, Argentina |
| 1985 - 1990 | MD, University of Buenos Aires, Buenos Aires, Buenos Aires, Argentina, Supervisor(s): NA |

## Professional Experience

| | |
|---|---|
| 2011 - present | Associate Professor in Medicine, Division of Neurology, Department of Medicine, St. Michael's Hospital-Unity Health. University of Toronto, Canada |
| 2009 - present | Scientist, Cardiovascular Program, Institute for Clinical Evaluative Sciences (ICES), Toronto, Ontario, Canada |
| 2007 - present | Scientist, Outcomes Research Unit, Li Ka Shing Knowledge Translation Institute, Toronto, Ontario, Canada |
| 2007 - present | Associate Professor, Institute of Health Policy, Management and Evaluation (iHPME), University of Toronto, Ontario, Canada |
| 2006 - present | Staff physician, Division of Neurology, Department of Medicine, St. Michael's Hospital, Toronto, Ontario, Canada |
| 1997 - 2003 | Staff physician, Department of Neurology, Ramos Mejia Hospital, University of Buenos Aires, Buenos Aires, Argentina |